

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
25. September 2003 (25.09.2003)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 03/078629 A1**

(51) Internationale Patentklassifikation<sup>7</sup>: **C12N 15/11**,  
15/82

(21) Internationales Aktenzeichen: PCT/EP03/02735

(22) Internationales Anmeldedatum:  
17. März 2003 (17.03.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
102 12 892.8 20. März 2002 (20.03.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von  
US): **BASF PLANT SCIENCE GMBH** [DE/DE]; Carl-  
Bosch-Strasse 38, 67056 Ludwigshafen (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **KOCK, Michael**  
[DE/DE]; Am Leutbusch 12, 67105 Schifferstadt (DE).  
**BAUER, Jörg** [DE/DE]; Friedrich-Profit-Str. 56, 67063  
Ludwigshafen (DE).

(74) Anwalt: **DÖRPER, Thomas**; BASF Aktiengesellschaft,  
67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,  
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE,  
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,  
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,  
MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO,  
RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ,  
UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE,  
DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL,  
PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG,  
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Veröffentlicht:**

- mit internationalem Recherchenbericht
- vor Ablauf der für Änderungen der Ansprüche geltenden  
Frist; Veröffentlichung wird wiederholt, falls Änderungen  
eintreffen

Zur Erklärung der Zweibuchstaben-Codes und der anderen  
Abkürzungen wird auf die Erklärungen ("Guidance Notes on  
Codes and Abbreviations") am Anfang jeder regulären Ausgabe  
der PCT-Gazette verwiesen.

(54) Title: CONSTRUCTS AND METHODS FOR THE REGULATION OF GENE EXPRESSION

(54) Bezeichnung: KONSTRUKTE UND VERFAHREN ZUR REGULATION DER GENEXPRESSION

(57) Abstract: The invention relates to constructs and methods for the regulation of gene expression of at least two endogenous target genes by introduction of an at least partly double-stranded ribonucleic acid molecule into a eukaryotic cell or a eukaryotic organism, whereby the ribonucleic acid molecule comprises at least two ribonucleotide sequence sections which are homologous with various genes of the eukaryotic cell.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft Konstrukte und Verfahren zur Regulation der Genexpression von mindestens zwei endogenen Zielgenen durch Einbringen eines zumindest teilweise doppelsträngigen Ribonukleinsäuremoleküls in eine eukaryotische Zelle oder einen eukaryotischen Organismus, wobei das Ribonukleinsäuremolekül mindestens zwei Ribonukleotidsequenzabschnitte umfasst, die zu verschiedenen Genen der eukaryotischen Zelle homolog sind.



**WO 03/078629 A1**

## Konstrukte und Verfahren zur Regulation der Genexpression

## Beschreibung

5

Die vorliegende Erfindung betrifft Konstrukte und Verfahren zur Regulation der Genexpression von mindestens zwei endogenen Zielgenen durch Einbringen eines zumindest teilweise doppelsträngigen Ribonukleinsäuremoleküls in eine eukaryotische Zelle oder einen eukaryotischen Organismus, wobei das Ribonukleinsäuremolekül mindestens zwei Ribonukleotidsequenzabschnitte umfasst, die zu verschiedenen Genen der eukaryotischen Zelle homolog sind.

Die gezielte Inhibition der Genexpression definierter Gene ist eine der am meisten beforschten Technologie der Biotechnologie. Die Expression von antisense-RNA ist dabei der am häufigsten verwendete Ansatz und vielfach beschrieben (u.a. EP-A1 0 458 367; EP-A1 0 140 308; van der Krol AR et al. (1988) BioTechniques 6(10):658-676; de Lange P et al. (1995) Curr Top Microbiol Immunol 197:57-75). Antisense-RNA vermittelte Ansätze haben jedoch den Nachteil, dass stöchiometrische Mengen der antisense-RNA erforderlich sind, um eine wirksame Inhibition der Ziel-mRNA zu bewirken. Weitere Probleme stehen im Zusammenhang mit dem Einbringen der antisense-RNA in ausreichenden Mengen in die Zellen und mit der Labilität der antisense-RNA. Ansätze basierend auf antisense-RNA sind daher meist ineffizient.

Ein weiterer Ansatz zur Genregulation ist die "Co-Suppression" und meint die Verminderung der Expression eines endogenen Zielgens durch transgene Expression einer sense-RNA dieses Zielgens (EP-A1 0 465 572). Der Co-Suppression liegen vermutlich mehr als ein Mechanismus zugrunde. Nachteilig ist die mangelnde Verlässlichkeit und Reproduzierbarkeit des Verfahrens. In manchen Fällen erfolgt Suppression, während in anderen Fällen - bedingt durch die Expression der sense-RNA - die erwartete Überexpression erfolgt. Auch ist der erhaltene Phänotyp oft nicht stabil. Die Anwendung der Co-Suppression ist im wesentlichen auf Pflanzen beschränkt.

Verschiedene Abwandlungen der Verfahren basierend auf antisense-RNA oder Cosuppression sind bekannt. So beschreibt WO 93/23551 ein Verfahren zur Inhibition mehrerer Gene durch Expression einer chimären antisense-RNA oder sense-RNA. Das Verfahren kann die üblichen mit antisense-RNA oder sense-RNA verbundenen Probleme nicht lösen und bleibt ineffizient.

WO 98/36083 und WO 99/15682 beschreiben die Regulation der Genexpression mittels viraler Expressionssysteme ("virus induced gene

silencing" VIGS).

WO 99/32619 und WO 99/53050 beschreiben Verfahren zur Inhibition einzelner Zielgene unter Verwendung einer RNA mit doppelsträngiger Struktur, wobei das Zielgen und die Region der RNA Duplex zu-  
5 mindest eine teilweise Identität aufweisen (siehe auch: Montgomery MK et al. (1998) Proc Natl Acad Sci USA 95:15502- 15507; Sharp PA (1999) Genes & Development 13(2):139-141; Fire A et al. (1998) Nature 391:806-11). Das Verfahren wird heute auch als  
10 "RNA-Interference" (RNAi) bezeichnet und hat in Mechanismus und Wirkung Ähnlichkeiten mit dem oben erwähnten VIGS Verfahren.

Die beschriebenen Verfahren, insbesondere das RNAi-Verfahren, lösen zwar einige Probleme im Zusammenhang mit der Verminderung  
15 einzelner Zielgene. Für andere Probleme, insbesondere für die parallele Suppression mehrerer Zielgene, konnte jedoch bislang keine befriedigende Lösung bereit gestellt werden. Zahlreiche Ansätze in der Biotechnologie erfordern nicht nur die Verminderung eines einzelnen Zielgens, sondern mehrerer Zielgene, wie bei-  
20 spielsweise verschiedener Gene eines oder verschiedener Stoffwechselwege oder ganzer Genfamilien. Bislang war dies nur mit erheblichen Arbeits- und Zeitaufwand zu realisieren. Die Ansätze erforderten oft die individuelle Regulation der einzelnen Zielgene durch sukzessive Transformation beispielsweise mit verschie-  
25 denen Expressionskonstrukten, die jeweils für eine antisense RNA eines Zielgens kodierten. Neben dem erheblichen Arbeits- und Zeitaufwand, besteht dabei der Nachteil, das für viele Systeme und Organismen nur eine beschränkte Anzahl von Selektionsmarkern, geeigneten Promotoren etc. zur Verfügung steht, was multiple  
30 Transformationen erheblich erschwert und beispielsweise die Deletion der Marker nach der Transformation und Selektion erfordert. Die mehrfache Verwendung eines Promotors hat oft unerwünschte Folgen, wie beispielsweise ein epigenetisches Gene-Silencing. Hierbei kommt es infolge der mehrfach verwendeten Kontrollsequenzen zu einer Inaktivierung derselben, vergleichbar der oben be-  
35 schriebenen Cosuppression.

Es stellte sich also die Aufgabe, neue Verfahren bereit zu stellen, die eine effiziente Verminderung der Expression mindestens  
40 zweier endogener Zielgene in einer eukaryotischen Zelle oder einem eukaryotischen Organismus ermöglichen. Diese Aufgabe wird durch die vorliegende Erfindung gelöst.

Ein erster Gegenstand der Erfindung betrifft Verfahren zur Verminderung der Expression von mindestens zwei verschiedenen, endo-  
45 genen Zielgenen in einer eukaryotischen Zelle oder einem eukaryotischen Organismus durch Einbringen eines zumindest teilweise doppelsträngigen Ribonukleinsäuremoleküls in besagte eukaryotische

## 3

Zelle oder besagten eukaryotischen Organismus, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

- 5 a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jeweils mindestens eine dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines jeden der besagten endogenen Zielgene und
- 10 b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.

Ein weiterer Gegenstand der Erfindung umfasst ein zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

- 20 a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jeweils mindestens eine dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens, wobei jedoch nicht alle "sense"-Ribonukleotidsequenzen zu dem "sense"-RNA-Transkript eines einzigen endogenen Zielgens identisch sind, und
- 25 b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.

30 Umfasst ist ferner die Verwendung der erfindungsgemäßen doppelsträngiges Ribonukleinsäuremolekül in einem der erfindungsgemäßen Verfahren.

Die vorliegende Erfindung löst die oben geschilderten Probleme und ermöglicht eine schnelle, besonders effiziente Methode zur Regulation der Expression verschiedener Zielgene. Insbesondere ergeben sich folgende Vorteile:

- 40 a) Transgene Organismen oder Zellen, in denen mehr als ein Zielgen inhibiert wird, können in einer einzigen Transformation erzeugt werden.
- 45 b) Die Transkriptionsrate für jeden Ribonukleotidsequenz der dsRNA ist gleich. Dadurch werden multiple Phänotypen durch unterschiedliche Expressionshöhen verhindert, wie sie bei individueller Expression separater Ribonukleotidsequenzen - beispielsweise durch den unterschiedlichen Ort der Insertion

## 4

in das Genom - oft entstehen. Dieser Vorteil gewährleistet eine gleichbleibend hohe Inhibition aller Zielgene und vermindert dramatisch die erforderlichen Selektionsschritte zu Generierung eines Organismus, bei dem alle Zielgene effizient  
5 supprimiert werden.

- c) Ein ökonomischer Umgang mit Kontrollelementen wie Promotoren und Selektionsmarkern wird ermöglicht. Zudem erübrigen sich Probleme, wie sie bei der mehrfachen Verwendung eines bestimmten Kontrollelementes, insbesondere eines Promoters,  
10 entstehen können ("epigenic gene silencing").
- d) Eine Segregation der einzelnen Ribonukleotidsequenzen bei nachfolgenden Züchtungs- und Kreuzungsschritten, wie sie bei der Verwendung mehrerer Expressionskonstrukte zwangsläufig  
15 entsteht, wird verhindert. Dadurch wird die nachfolgende Züchtung stabiler Linien erheblich erleichtert und beschleunigt.
- e) Organismen mit komplexen beispielsweise polyploide Genomen, wie beispielsweise manche Pflanzen, sind einer effizienten Gensuppression zugänglich. Aufgrund der zahlreichen Kopien für einzelne Gene sind diese Organismen klassischen verfahren  
20 der Mutagenese und Selektion nicht zugänglich.

25 Überraschenderweise konnte bei dem erfindungsgemäßen Verfahren keine störende Interferenz zwischen den einzelnen Ribonukleotidsequenzabschnitte untereinander beobachtet werden.

- 30 "Endogenes Zielgen einer eukaryotischen Zelle oder eines eukaryotischen Organismus" meint jede Nukleinsäuresequenz in einer eukaryotischen Zelle, einem eukaryotischen Organismus oder einem Teil, Organ, Gewebe, Samen etc. desselben, die zur Transkription befähigt ist. Dabei kann es sich um natürlicherweise vorkommende oder  
35 aber künstlich eingeführte Sequenzen (wie beispielsweise transgene Sequenzen) handeln, wobei natürlicherweise vorkommende Sequenzen bevorzugt sind. Natürlicherweise vorkommende Sequenzen sind bevorzugt und umfassen sowohl die eigenen Sequenzen der eukaryotischen Zelle oder des eukaryotischen Organismus als auch  
40 Gene von Pathogenen, die in der eukaryotischen Zelle oder dem eukaryotischen Organismus nach einem Befall durch ein Pathogen präsent sind. Das Zielgen kann in der chromosomalen DNA oder der DNA der Organellen (wie beispielsweise der Plastiden z.B. Chloroplasten etc.) lokalisiert sein oder aber sich extrachromosomal in  
45 der Zelle befinden. Die natürlicherweise vorkommenden, eigenen Sequenzen des eukaryotischen Organismus umfassen bevorzugt Gene desselben, die stabil im Genom vorliegen, wobei das Genom die Ge-

## 5

samtheit der genetischen Information meint und sowohl die chromosomale als auch die plastidäre DNA umfasst. Bevorzugt ist das endogene Zielgen ein natürlicherweise in der chromosomalen DNA vorkommendes Gen. Bevorzugt sind Gene deren verminderte Expression zu einem veränderten Phänotyp führt.

"Verminderung" oder "vermindern" der Expression eines Zielgens ist im Zusammenhang weit auszulegen und umfasst die teilweise oder im wesentlichen vollständige, auf unterschiedliche zellbiologische Mechanismen beruhende Unterbindung oder Blockierung der Expression des Zielgens oder der von ihm abgeleiteten RNA, mRNA, rRNA, tRNA und/oder des dadurch kodierten Proteinproduktes in einer Zelle oder einem Organismus oder einem davon abgeleiteten Teil, Gewebe, Organ, Zelle oder Samen. Eine Verminderung im Sinne der Erfindung umfasst die mengenmäßige Verringerung einer vom Zielgen exprimierten RNA, mRNA, rRNA, tRNA und/oder des dadurch kodierten Proteinproduktes bis hin zu einem im wesentlichen vollständigen Fehlen derselben. Dabei wird die Expression einer bestimmten RNA, mRNA, rRNA, tRNA und/oder des dadurch kodierten Proteinproduktes in einer Zelle oder einem Organismus im Vergleich zu der selben Zelle oder Organismus, die dem Verfahren nicht unterworfen wurden, bevorzugt um mehr als 50%, besonders bevorzugt um mehr als 80%, ganz besonders bevorzugt um mehr als 90%, am meisten bevorzugt mehr als 95% vermindert. Dabei kann die Verminderung durch dem Fachmann geläufigen Verfahren ermittelt werden. So kann die Verminderung der Proteinmenge beispielsweise durch immunologischen Nachweis des Proteins bestimmt werden. Weiterhin können biochemische Techniken wie Northern-Hybridisierung, "nuclease protection assay", Reverse Transkription (quantitative RT-PCR), ELISA ("enzyme linked immunosorbent assay"), Western-Blotting, Radioimmunoassay (RIA) oder andere Immunoassays sowie "fluorescence activated cell analysis" (FACS) eingesetzt werden. Je nach Art des verminderten Proteinproduktes kann auch dessen Aktivität oder der Einfluss auf den Phänotyp des Organismus oder der Zelle ermittelt werden.

"Proteinmenge" meinte die Menge eines bestimmten Polypeptides in einem Organismus, einem Gewebe, einer Zelle oder einem Zellkompartiment.

"Verminderung" der Proteinmenge meint die mengenmäßige Verminderung der Menge eines bestimmten Polypeptides in einem Organismus, einem Gewebe, einer Zelle oder einem Zellkompartiment - beispielsweise durch das erfindungsgemäße Verfahren - im Vergleich zu dem Wildtyp derselben Gattung und Art auf den dieses Verfahren nicht angewendet wurde, unter ansonst gleichen Rahmenbedingungen (wie beispielsweise Kulturbedingungen, Alter, Nährstoffzufuhr

etc.). Der Verminderung beträgt dabei mindestens 10 %, bevorzugt mindestens 10% oder mindestens 20%, besonders bevorzugt um mindestens 40% oder 60%, ganz besonders bevorzugt um mindestens 70% oder 80%, am meisten bevorzugt um mindestens 90% oder 95%. Verfahren zur Bestimmung der Proteinmenge sind dem Fachmann bekannt. Beispielfhaft seien zu nennen: Das Mikro-Biuret Verfahren (Goa J (1953) Scand J Clin Lab Invest 5:218-222), die Folin-Ciocalteu-Methode (Lowry OH et al. (1951) J Biol Chem 193:265-275) oder die Messung der Adsorption von CBB G-250 (Bradford MM (1976) Analyst Biochem 72:248-254).

"Verschieden" meint in Bezug auf zwei endogene Zielgene bevorzugt, dass die von den beiden endogenen Zielgenen transkribierte RNA oder mRNA nicht identisch ist. Bevorzugt ist die Homologie der von den beiden endogenen Zielgenen transkribierte RNA oder mRNA geringer als 90%, bevorzugt geringer als 80%, besonders bevorzugt geringer als 70%, ganz besonders bevorzugt geringer als 60%, am meisten bevorzugt geringer als 50% über jeweils die gesamte Länge der transkribierten RNA oder mRNA.

"Zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül" (infolge dsRNA) meint Ribonukleinsäuremolekül, die ganz oder teilweise doppelsträngig sind. Bevorzugt ist die Ribonukleinsäuresequenz überwiegend vollständig doppelsträngig. "Überwiegend vollständig doppelsträngig" meint, dass zumindest 50%, bevorzugt 70%, besonders bevorzugt 80%, ganz besonders bevorzugt 90% der in dem Molekül vorhandenen Basen in Paarung mit einer anderen Base der dsRNA vorliegen oder - entsprechend der Sequenz der dsRNA und den Basenpaarregeln sowie gegebenenfalls einer RNA-Sekundärstrukturvoraussage mittels eines geeigneten Computeralgorithmus - zumindest theoretisch vorliegen können.

"Im wesentlichen identisch" meint, dass eine "sense"-Ribonukleotidsequenz der dsRNA auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu der Sequenz des "sense"-RNA-Transkriptes eines endogenen Zielgens aufweisen kann. Mutationen umfassen Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Basen einer Nukleinsäuresequenz. Bevorzugt beträgt die Homologie zwischen einer "sense"-Ribonukleotidsequenz einer dsRNA und mindestens einem Teil des "sense"-RNA-Transkript eines endogenen Zielgens mindestens 60 %, bevorzugt mindestens 70 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt 95%. Die Sequenzen können auch identisch mit der korrespondierenden Sequenz des Zielgens sein. Eine 100%ige Sequenzidentität zwischen der "sense"-Ribonukleotidsequenz der dsRNA und mindestens einem Teil des "sense"-Stranges der Transkriptes eines endogenen Gens ist bevorzugt, wenn gleich

nicht zwingend erforderlich, um eine effiziente Verminderung der Expression des endogenen Gens zu bewirken. Einzelne Mutationen werden toleriert. Das Verfahren ist demnach tolerant gegenüber Sequenzabweichungen, wie sie infolge genetischer Mutationen, Polymorphismen oder evolutionärer Divergenzen vorliegen können. So ist es beispielsweise auch möglich mit einer einzigen dsRNA, die ausgehend von einer bestimmten endogenen Gen generiert wurde, die Expression weiterer homologer endogener Gene des gleichen Organismus oder aber auch die Expression homologer endogener Gene in anderen verwandten Arten zu unterdrücken.

Unter Homologie wird das Maß der Übereinstimmung zwischen zwei Nukleotid-, Ribonukleotid- oder Proteinsequenzen verstanden, die bevorzugt durch Vergleich mit Hilfe des Programmalgorithmus GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA; Altschul et al. (1997) Nucleic Acids Res. 25:3389ff) unter Einstellung folgender Parameter berechnet wird:

20 Gap Weight: 50

Length Weight: 3

Average Match: 10

Average Mismatch: 0

Dem Fachmann ist bewusst, dass wenn die Homologie zwischen DNA (z.B. Genen) und RNA bestimmt wird, Thymin (T) in der DNA Sequenz als äquivalent zu Uracil (U) in der RNA Sequenz betrachtet wird.

"Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens" meint Fragmente einer RNA oder mRNA transkribiert von einem endogenen Zielgen. Dabei hat besagtes Teil bevorzugt eine Sequenzlänge von mindestens 10 Basen, bevorzugt mindestens 25 Basen, besonders bevorzugt mindestens 50 Basen, ganz besonders bevorzugt mindestens 100 Basen, am meisten bevorzugt mindestens 200 Basen oder mindestens 300 Basen. Umfasst ist auch die vollständige transkribierte RNA oder mRNA.

Alternativ, kann eine "im wesentlichen identische" dsRNA auch als Nukleinsäuresequenz definiert werden, die befähigt ist, mit einem Teil eines Transkriptes, bevorzugt der mRNA, eines endogenen Zielgenes zu hybridisieren (z.B. in 400 mM NaCl, 40 mM PIPES pH 6,4, 1 mM EDTA bei 50°C oder 70°C für 12 bis 16 h oder unter anderen Standardhybridisierungsbedingungen).

"Standardhybridisierungsbedingungen" ist breit zu verstehen und meint weniger stringente als auch - bevorzugt - stringente Hybridisierungsbedingungen. Solche Hybridisierungsbedingungen sind unter anderem bei Sambrook J, Fritsch EF, Maniatis T et al., in



## 8

Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57) oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

5

Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit ungefähr 2X SSC bei 50°C) und - bevorzugt - solchen mit hoher Stringenz (mit ungefähr 0,2X SSC bei 50°C bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M NaCl, pH 7.0). Darüberhinaus kann die Temperatur während des Waschschrilles von niedrig stringenten Bedingungen bei Raumtemperatur, ungefähr 22°C, bis zu - bevorzugt - stärker stringenten Bedingungen bei ungefähr 65°C angehoben werden. Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50% Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt. Einige beispielhafte Bedingungen für Hybridisierung und Waschschrill sind in-

(1) Hybridisierungsbedingungen zum Beispiel aus nachfolgenden Bedingungen ausgewählt sein:

- a) 4X SSC bei 65°C,
- b) 6X SSC bei 45°C,
- c) 6X SSC, 100 µg/ml denaturierter, fragmentierte Fischsperma-DNA bei 68°C,
- f) 50% Formamid, 4X SSC bei 42°C,
- h) 2X oder 4X SSC bei 50°C (schwach stringente Bedingung),
- i) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (schwach stringente Bedingung).

35

(2) Waschschrill können zum Beispiel aus nachfolgenden Bedingungen ausgewählt sein:

- a) 0,015 M NaCl/0,0015 M Natriumcitrat/0,1% SDS bei 50°C.
- b) 0,1X SSC bei 65°C.
- c) 0,1X SSC, 0,5% SDS bei 68°C.
- d) 0,1X SSC, 0,5% SDS, 50% Formamid bei 42°C.
- e) 0,2X SSC, 0,1% SDS bei 42°C.
- f) 2X SSC bei 65°C (schwach stringente Bedingung).

45

"Im wesentlichen komplementär" meint, dass die "antisense"-Ribonukleotidsequenzen der dsRNA auch Insertionen, Deletionen sowie

- einzelne Punktmutationen im Vergleich zu dem Komplement der "sense"-Ribonukleotidsequenzen aufweisen kann. Bevorzugt beträgt die Homologie mindestens 80 %, bevorzugt mindestens 90 %, ganz besonders bevorzugt mindestens 95 %, am meisten bevorzugt 100%
- 5 zwischen den "antisense"-Ribonukleotidsequenzen und dem Komplement der "sense"-Ribonukleotidsequenzen. Komplement meint dabei - in der dem Fachmann geläufigen Weise - den entsprechend den Basenpaarregeln abgeleiteten Gegenstrang.
- 10 Die doppelsträngige Struktur der dsRNA kann ausgehend von einem einzigen, ganz oder teilweise selbstkomplementären RNA-Strang (bei dem die oben erwähnten "sense"- und "antisense"-Ribonukleotidsequenzen der dsRNA alle kovalent miteinander verbunden sind) oder ausgehend von zwei RNA-Strängen (indem die oben erwähnten
- 15 "sense"- und "antisense"-Ribonukleotidsequenzen der dsRNA auf separate Stränge vorliegen), die zueinander ganz oder teilweise komplementär sind, gebildet werden. Bei zwei separaten Strängen können beispielsweise alle "sense"-Ribonukleotidsequenzen auf dem einen und alle "antisense"-Ribonukleotidsequenzen auf dem anderen
- 20 Strang vorliegen. Die Sequenzen können aber auch anders auf die beiden Stränge verteilt sein. Die Ausbildung der doppelsträngigen Struktur kann in vitro aber auch in vivo - beispielsweise in der eukaryotischen Zelle selber - erfolgen. Bevorzugt liegt die dsRNA in Form eines einzigen, selbstkomplementären RNA-Stranges vor.
- 25 Die einzelnen "sense"-Ribonukleotidsequenzen können mit den korrespondierenden, im wesentlichen komplementären "antisense"-Ribonukleotidsequenzen eine doppelsträngige RNA-Struktur mittels Basenpaarung ausbilden und bilden eine Untereinheit der dsRNA.
- 30 Im Falle eines selbstkomplementären Stranges ergeben sich verschiedene Möglichkeiten für die Primärstruktur der dsRNA. Nachfolgend aufgeführte sind beispielhaft, jedoch nicht einschränkend zu verstehen:
- 35 a) Es können zunächst die "sense"-Ribonukleotidsequenzen (S) der einzelnen Untereinheiten aneinander gefügt werden, worauf dann eine Aneinanderreihung der im wesentlichen komplementären "antisense"-Ribonukleotidsequenzen (AS) folgt. Die Anzahl der
- 40 Einheiten n ist größer oder gleich zwei. Es entsteht eine Struktur mit einer einzelnen Haarnadel. Die Primärstruktur der dsRNA kann dabei schematisch beispielsweise wie folgt aussehen:
- 45 5'-S(1)-S(2)-.....-S(n)-AS(n)-....-AS(2)-AS(1)-3'

## 10

Die bevorzugte Sekundärstruktur ist in Fig. 2-A wiedergegeben.

- b) Es können zunächst die "sense"-Ribonukleotidsequenz (S) und die im wesentlichen komplementäre "antisense"-Ribonukleotidsequenz (AS) der ersten Untereinheiten aneinander gefügt werden, wodauf dann die Aneinanderreihung von "sense"- und "antisense"-Ribonukleotidsequenzen der weiteren Untereinheiten folgt. Die Anzahl der Einheiten n ist größer oder gleich zwei. Es entsteht eine Struktur mit mehreren Haarnadeln. Die Primärstruktur der dsRNA kann dabei schematisch beispielsweise wie folgt aussehen:

5'-S(1)-AS(1)-S(2)-AS(2).....-S(n)-AS(n)-3'

15

Die bevorzugte Sekundärstruktur ist in Fig. 2-B wiedergegeben.

- Ist die dsRNA - bevorzugt - in der Lage eine Haarnadelstruktur auszubilden, so entspricht der Stamm der Haarnadel dem doppelsträngige Anteil der dsRNA, der durch Basenpaarung zwischen auf dem gleich RNA-Moleküle lokalisierten "sense"- und "antisense"-Ribonukleotidsequenz gebildet wird. Dabei werden "sense"- und "antisense"-Ribonukleotidsequenzen bevorzugt durch einen "Linker" verbunden. Der "Linker" ist bevorzugt ein Intron, das aus der dsRNA herausgespleißt werden kann. Selbstkomplementären dsRNA-Strukturen ausgehend von einem einzelnen RNA-Molekül sind bevorzugt, da sie lediglich die Expression eines Konstruktes erfordern und die komplementären RNA-Stränge stets in einem äquimolaren Verhältnis umfassen.

30

Bei der Verwendung eines Linkers (I) - bevorzugt eines Intron - seien nachfolgende schematische Primärstrukturen für die dsRNA beispielhaft genannt:

- 35 c) Dies ist eine bevorzugte Variante von a), bei der an der Stelle der Haarnadelschlaufe ein Linker (I) - bevorzugt ein Intron - insertiert wird:

5'-S(1)-S(2)-.....-S(n)-I-AS(n)-....-AS(2)-AS(1)-3'

40

Die bevorzugte Sekundärstruktur ist in Fig. 2-C wiedergegeben.

- d) Dies ist eine bevorzugte Variante von b), bei der an der Stelle der jeder Haarnadelschlaufe ein Linker (I) - bevorzugt ein Intron - insertiert wird:

5'-S(1)-I-AS(1)-S(2)-I-AS(2).....S(n)-I-AS(n)-3'

Die bevorzugte Sekundärstruktur ist in Fig. 2-D wiedergegeben.  
5      ben.

Die dsRNA Moleküle sind jedoch auch ohne den Linker funktionell. Dabei ist jedoch zu berücksichtigen, dass die letzten ca. 10 Nukleotide der terminalen Untereinheit S(n) in diesem Fall nicht  
10 mehr korrekt paaren. In diesem Fall ist die Länge für diese Untereinheit um 10 Nukleotide zu ergänzen. Der Linker ist bevorzugt ein Intron, besonders bevorzugt ein Intron in sense-Orientierung. Bevorzugt handelt es sich um ein Intron eines pflanzlichen Gens. Beispielfhaft jedoch nicht einschränkend seien zu nennen: Das In-  
15 tron 3 der Alkoholdehydrogenase 1' (Adh1) aus Mais (GenBank Acc.-No.: AF044293; GI: 2828164), das Intron 4 der beta-Conglycinin alpha Untereinheit aus Soja (GenBank Acc.-No.: AB051865); eines der Introns des rbcS-3A Gens für Ribulose-1.5-bisphosphatcarboxylase (RBC) kleine Untereinheit aus Erbse (GenBank Acc.-No.:  
20 X04333). Diese und weitere geeignete Introns sind dem Fachmann bekannt (McCullough AJ & Schuler MA (1997) Nuc Acids Res 25:1071-1077). Für die Anwendung in dem erfindungsgemäßen Verfahren wird das Intron bevorzugt in Kombination mit Spleißakzeptor- und Spleißdonorsequenzen eingesetzt, die ein späteres Heraus-  
25 spleißen aus der dsRNA ermöglichen. Diese Spleißsequenzen können die flankierenden Sequenzen des Intron selber sein, oder aber auch durch entsprechende Sequenzen der übrigen dsRNA bereitgestellt werden.

30 Jede der einzelnen "sense"-Ribonukleotidsequenzen der dsRNA ist im wesentlichen identisch zu mindestens einem Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens. Dabei sind jedoch nicht alle "sense"-Ribonukleotidsequenzen zu dem "sense"-RNA-Transkript eines einzigen endogenen Zielgens iden-  
35 tisch, sondern die jeweils maximale Identität von mindestens zwei der "sense"-Ribonukleotidsequenzen besteht zu den "sense"-RNA-Transkripten von unterschiedlichen endogenen Zielgenen. Dabei beträgt die Homologie zwischen den Transkripten der beiden endogenen Zielgene unter 90%, bevorzugt unter 80%, besonders bevorzugt  
40 unter 70%, ganz besonders bevorzugt unter 60%, am meisten bevorzugt unter 50%.

Mindestens zwei der in der erfindungsgemäßen dsRNA umfassten einzelnen "sense"-Ribonukleotidsequenzen sind unterschiedlich. Un-  
45 terschiedlich bedeutet zum einen, dass die Zielgene zu deren Transkripten sie die jeweils maximale Identität aufweisen, nicht identisch sind. Bevorzugt vermindert mindestens eine Untereinheit

## 12

der dsRNA die Expression eines anderen Gens als mindestens eine andere Untereinheit. Unterschiedlich kann zum anderen auch heißen, dass die "sense"-Ribonukleotidsequenzen der Untereinheiten selber im wesentlichen nicht identisch sind und bevorzugt eine  
5 Homologie zu einander unter 60%, besonders bevorzugt unter 50% ganz besonders bevorzugt unter 40% aufweisen. Die dsRNA kann in einer weiteren Ausführungsform mehrerer Kopien einer Untereinheit enthalten. Weiterhin kann die dsRNA auch mehrere verschiedene Untereinheiten enthalten, die aber gegen das gleiche endogene Ziel-  
10 gens gerichtet sind und deren "sense"-Ribonukleotidsequenzen beispielsweise im wesentlichen identisch sind zu unterschiedlichen Teilen des "sense"-RNA-Transkriptes des besagten endogenen Zielgens.

15 Dabei kann jede der einzelnen "sense"-Ribonukleotidsequenzen auch zu dem Transkript mehrerer endogener Zielgene im wesentlichen identisch sein. Dies ist besonders dann der Fall, wenn die Zielgene über ähnliche Sequenzabschnitte verfügen, wie es beispielsweise bei Mitgliedern von Genfamilien (z.B. Speicherproteinen)  
20 der Fall ist. Dies ist eine besonders vorteilhafte Anwendungsform, da - bei entsprechender Wahl der Ribonukleotidsequenz einer Untereinheit - besagte Untereinheit die Expression von mehr als einem Zielgen vermindern kann.

25 Vorzugsweise wird die Sequenz der dsRNA so gewählt, dass die angestrebte dsRNA Struktur nach Ausbildung der Duplex - im Vergleich zu anderen möglichen Faltungsvarianten der Primärstruktur der dsRNA - die jeweils geringste freie Energie hat. Dies kann beispielsweise durch Vermeidung von Sequenzduplikationen etc. ge-  
30 währleistet werden. Die spezifische Sekundärstruktur kann beispielsweise mit geeigneten Computerprogrammen vorausgesagt und optimiert werden (z.B. FOLDRNA; Zuker and Stiegler (1981) Nucleic Acids Res 9(1):133-48).

35 Jede Untereinheit der dsRNA hat in einer bevorzugten Ausführungsform eine Länge von mindestens 20 Basenpaaren, bevorzugt mindestens 50 Basenpaaren, besonders bevorzugt mindestens 100 Basenpaare, ganz besonders bevorzugt mindestens 250 Basenpaare.

40 In einer weiterhin bevorzugten Ausführungsform hat jede Einheit eine Länge eine ganzzahligen Vielfachen von 21 oder 22 Basenpaaren, also beispielsweise 21, 22, 42, 43, 44, 63, 64, 65, 66, 84, 85, 86, 87, 88, 105, 106, 107, 108, 109, 110, 126, 127, 128, 129, 131, 132, 147, 148, 149, 150, 151, 152, 153, 154, 168, 169, 170,  
45 171, 172, 173, 174, 175, 176, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219 oder 220 Basenpaare, bevorzugt 21, 22, 42, 44, 63, 66, 84, 88,

## 13

105, 110, 126, 132, 147, 154, 168, 176, 189, 198, 210 oder 220 Basenpaare, ganz besonders bevorzugt 21, 42, 63, 84, 105, 126, 147, 168, 189 oder 210 Basenpaare, am meisten bevorzugt 180 oder 210 Basenpaare.

5

Die "sense"- und/oder "antisense"-Ribonukleotidsequenzen der einzelnen Untereinheiten können direkt oder aber durch einen "Spacer" (SP; Abstandshalter) miteinander verbunden und/oder flankiert sein. Die einzelnen "Spacer" (SP) können dabei identisch  
 10 oder aber auch unterschiedlich sein. Der "Spacer" genügt dabei bevorzugt den gleichen Längenanforderungen wie sie oben für die Länge der Untereinheiten selber gegeben sind. Der "Spacer" kann eine doppelstränge Struktur ausbilden, kann aber auch - beispielsweise in Form einer Blase - in ungepaarter Formation bestehen, d.h. die Basen in Strang und Gegenstrang müssen nicht zwingenderweise komplementär sein. Bevorzugte Ausführungsformen sind  
 15 zum Beispiel durch nachfolgende Primärstrukturen beschrieben:

e) Dies ist eine bevorzugte Variante von c):

20

5' SP-S(1)-SP-S(2)-SP-...-S(n)-AS(n)-SP-...-AS(2)-SP-AS(1)-SP-3'

Die bevorzugte Sekundärstruktur ist in Fig. 3-A wiedergegeben.

25

Der "Spacer" kann weitere Funktionselemente umfassen. Beispielshaft jedoch nicht einschränkend sind zu nennen:

i) Sequenzen kodierend für eine von einem Ribozym als Substrat  
 30 erkannten Erkennungssequenz (RE). Beispielsweise kann die dsRNA nachfolgende lineare Struktur vor der Faltung einnehmen:

5'-S(1)-(RE)-S(2)-...-S(n)-AS(n)-...-AS(2)-(RE)-AS(1)-3'

35

Die bevorzugte Sekundärstruktur ist in Fig. 3-B wiedergegeben. Das entsprechende Ribozym (R) kann separat exprimiert werden kann aber auch auf der dsRNA selber kodiert sein. Dabei ist die Sequenz kodierend für ein Ribozym bevorzugt so  
 40 angeordnet, dass sie im gefalteten dsRNA Molekül einer Sequenz gegenüber liegt, die für dieses Ribozym als Substrat fungieren kann. Beispielsweise kann die dsRNA nachfolgende lineare Struktur vor der Faltung einnehmen:

45 5'-S(1)-(R)(RE)-S(2)-...-S(n)-AS(n)-...-AS(2)-(R)(RE)-AS(1)-3'

## 14

Die bevorzugte Sekundärstruktur ist in Fig. 3-C wiedergegeben. Durch die genannten Ausführungsformen werden nach Transkription die einzelnen Untereinheiten durch Wirkung des Ribozym voneinander getrennt. Diese Trennung ist vorteilhaft,  
5 jedoch nicht zwingend erforderlich. Entsprechend nutzbare Ribozyme und Erkennungssequenzen sind dem Fachmann bekannt.

Ribozyme meint katalytische RNA-Moleküle. Ribozyme können an jede beliebige Ziel-RNA angepasst werden und spalten das  
10 Phosphodiester-Gerüst an spezifischen Positionen, wodurch die Ziel-RNA funktionell deaktiviert wird (Tanner NK (1999) FEMS Microbiol Rev 23(3):257-275). Das Ribozym wird dadurch nicht selber modifiziert, sondern ist in der Lage, weitere Ziel-RNA-Moleküle analog zu spalten, wodurch es die Eigenschaften  
15 eines Enzyms erhält. Der Einbau von Ribozymsequenzen in "antisense"-RNAs verleiht eben diesen "antisense"-RNAs diese enzymähnliche, RNA-spaltende Eigenschaft und steigert so deren Effizienz bei der Inaktivierung der Ziel-RNA. Die Herstellung und Verwendung entsprechender Ribozym-"antisense"-RNA-Moleküle ist beispielsweise beschrieben bei Haseloff et al.  
20 (1988) Nature 334: 585-591. Auf diese Art können Ribozyme (z.B. "Hammerhead"-Ribozyme; Haselhoff und Gerlach (1988) Nature 334:585-591) verwendet werden, um die eines bestimmte RNA katalytisch zu spalten. Verfahren zur Expression von Ribozymen zur Verminderung bestimmter Proteine sind beschrieben  
25 in (EP 0 291 533, EP 0 321 201, EP 0 360 257). In pflanzlichen Zellen ist eine Ribozym-Expression ebenfalls beschrieben (Steinecke P et al. (1992) EMBO J 11(4):1525-1530; de Feyter R et al. (1996) Mol Gen Genet. 250(3):329-338). Geeignete  
30 Zielsequenzen und Ribozyme können zum Beispiel wie bei "Steinecke P, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds, Academic Press, Inc. (1995), S.449-460" beschrieben, durch Sekundärstrukturberechnungen von Ribozym- und Ziel-RNA sowie durch deren Interaktion bestimmt werden (Bayley CC et  
35 al. (1992) Plant Mol Biol. 18(2):353-361; Lloyd AM and Davis RW et al. (1994) Mol Gen Genet. 242(6):653-657). Beispielsweise können Derivate der Tetrahymena L-19 IVS RNA konstruiert werden, die komplementäre Bereiche zu der mRNA des  
40 zu den Spacersequenzen aufweisen (siehe auch US 4,987,071 und US 5,116,742). Alternativ können solche Ribozyme auch über einen Selektionsprozess aus einer Bibliothek diverser Ribozyme identifiziert werden (Bartel D und Szostak JW (1993) Science 261:1411-1418).

45 ii) Sequenzen kodierend für Erkennungssequenzen für RNAasen  
Der "Spacer" kann Erkennungssequenzen für RNAasen, bevorzugt sequenzspezifische RNAasen wie beispielsweise RNase III ent-

## 15

halten. RNase III schneidet am Motiv 5'-AGNN-3, wenn vier dieser Motive in einer Schleife vorhanden sind (Nagel R & Ares M (2000) RNA 6:1142-1156). Die RNase kann eine pflanzeneigene RNase sein, oder - wie beispielsweise für bakterielle RNase III Proteine - auch transgen exprimiert werden.

iii) Sequenzen kodierend für Intronspleißsignale (IS). Dabei sind die Spleißdonor und Spleißakzeptorsequenzen bevorzugt so lokalisiert, dass jeweils die Untereinheit als Intron herausgespleißt wird. Intronspleißsignale sind in Meritt et al. (1997) Plant Journal 12:937-943 oder in Egoavil et al. (1997) Plant Journal 12:971-980 beschrieben.

Die dsRNA bzw. ihre Vorläufermoleküle können auf verschiedene dem Fachmann geläufige Weise in einen Organismus oder eine Zelle eingebracht werden. "Einbringen" ist breit zu verstehen und umfasst im Rahmen der Erfindung alle Verfahren, die dazu geeignet eine dsRNA bzw. ihre Vorläufermoleküle, direkt oder indirekt, in einen Organismus oder eine Zelle, Kompartiment, Gewebe, Organ oder Samen desselben einzuführen oder dort zu generieren. Die Einbringung kann zu einer vorübergehenden (transienten) Präsenz einer dsRNA führen oder aber auch zu einer dauerhaften (stabilen). Umfasst sind Verfahren der direkten Transfektion oder Transformation der Zelle mit der als auch die Transformation oder Transfektion der Zelle mit Expressionskassetten, die befähigt sind, die der dsRNA zugrundeliegenden Ribonukleinsäuresequenzen in der Zelle zu exprimieren (infolge dsRNA-Expressionssystem). Die Expression der dsRNA kann transient oder - beispielsweise nach Integration in das Genom des Organismus - permanent erfolgen. Die Duplex-Bildung der dsRNA kann entweder außerhalb der Zelle oder innerhalb derselben initiiert werden.

Die dsRNA wird in einer Menge eingeführt, die zumindest eine Kopie pro Zelle ermöglicht. Höhere Mengen (z.B. mindestens 5, 10, 100, 500 oder 1000 Kopien pro Zelle) können ggf. eine effizientere Verminderung der Expression der Zielgene bewirken. Da dsRNA eine außerordentlich gute Mobilität innerhalb eines Organismus hat, ist es nicht zwingend erforderlich die dsRNA in jede Zelle des Organismus zu applizieren. Es ist ausreichend, die dsRNA in eine oder wenige Zellen einzubringen oder zu exprimieren, wobei die erfindungsgemäße Wirkung dann auch in anderen Zellen des gleichen Organismus erzielt werden kann.

Eine dsRNA - beispielsweise zur Verwendung in einer direkten Transformation oder Transfektion - kann in vivo oder in vitro, durch enzymatische, molekularbiologische oder chemisch-synthetische Verfahren synthetisiert werden. Dazu können eukaryoti-



- sche, prokaryotische oder Bakteriophagen RNA Polymerasen (wie z.B. T3-, T7- oder SP6 RNA-Polymerase) verwendet werden. Entsprechende Verfahren zu in vitro Expression von RNA sind beschrieben (WO 97/32016; US 5,593,874; US 5,698,425, US 5,712,135, US 5,789,214, US 5,804,693). Eine chemisch oder enzymatisch in vitro synthetisierte dsRNA kann vor der Einführung in eine Zelle, Gewebe oder Organismus aus dem Reaktionsgemisch beispielsweise durch Extraktion, Präzipitation, Elektrophorese, Chromatographie oder Kombinationen dieser Verfahren ganz oder teilweise aufgereinigt werden. Die dsRNA kann direkt in die Zelle eingeführt werden (beispielsweise durch Partikelbeschuss oder Mikroinjektion) oder aber extrazellulär (z.B. in den interstitial Raum, das Gefäßsystem, das Verdauungssystem o.ä.) appliziert werden. Auch eine Applikation beispielsweise von dsRNA exprimierenden Organismen in Form von Nahrung ist denkbar. Es ist bekannt, dass dsRNA eine gute Zellgängigkeit und ausreichende Stabilität hat. Durch die hohe Wirksamkeit der dsRNA sind auch wenige Moleküle ausreichend, um eine gute Wirkung im Sinne der Erfindung zu erzielen.
- 20 Es können ferner Modifikationen sowohl des Zucker-Phosphat-Gerüsts als auch der Nukleoside in der dsRNA vorliegen. Beispielsweise können die Phosphodiesterbindungen der RNA dahingehend modifiziert sein, dass sie zumindest ein Stickstoff oder Schwefel-Heteroatom umfassen. Basen können dahingehend modifiziert werden, dass die Aktivität beispielsweise von Adenosindeaminase eingeschränkt wird. Die dsRNA kann enzymatisch oder ganz oder teilweise chemisch-synthetisch hergestellt werden.

- Bevorzugt wird die dsRNA jedoch ausgehend von entsprechenden Expressionssystemen in der Zelle exprimiert. Ein weiterer Gegenstand der Erfindung betrifft besagte dsRNA-Expressionssysteme. Wird die dsRNA als ein einzelner, selbstkomplementärer RNA-Strang exprimiert, so umfasst das Expressionssystem eine Expressionskassette mit einer für den selbstkomplementären RNA-Strang kodierenden DNA Sequenz in funktioneller Verknüpfung mit einem Promotor, der geeignet ist, die Expression in der jeweiligen eukaryotischen Zelle zu gewährleisten. Optional kann die Expressionskassette weitere funktionelle Elemente wie beispielsweise Transkriptionsterminatoren und/oder Polyadenylierungssignale umfassen. Derartige Expressionskassetten sind ebenfalls Gegenstand der Erfindung.

- Wird die dsRNA in Form von zwei separaten Strängen exprimiert, die zueinander ganz oder teilweise komplementär sind, so umfasst das Expressionssystem zwei Expressionskassetten, wobei jeder der beiden Stränge in funktioneller Verknüpfung mit einem Promotor steht, der geeignet ist, die Expression in der jeweiligen euka-

## 17

ryotischen Zelle zu gewährleisten. Optional können die Expressionskassetten weitere funktionelle Elemente wie beispielsweise Transkriptionsterminatoren und/oder Polyadenylierungssignale umfassen. Die Kombination der beiden Expressionskassetten zu dem  
5 erfindungsgemäßen Expressionssystem kann auf verschiedene dem Fachmann geläufige Art geschehen. Beispielhaft seien zu nennen:

- a) Transformation der Zelle oder Pflanze mit einem Vektor, der Expressionskassetten für beide RNA-Stränge umfasst,  
10
- b) Kotransformation der Zelle oder Pflanze mit zwei Vektoren, wobei jeweils ein Vektor für jeweils einen der beiden Stränge der dsRNA kodiert.
- 15 c) Kreuzung von zwei Pflanzen, die mit jeweils einem Vektor transformiert wurden, wobei jeweils ein Vektor für jeweils einen der beiden Stränge der dsRNA kodiert.

Es ist auch möglich, das eine Expressionskassette einzusetzen,  
20 bei der die für die dsRNA kodierende DNA-Sequenz zwischen zwei Promotoren mit entgegengerichteter Transkriptionsrichtung lokalisiert ist und so von beiden Seiten transkribiert wird.

Expressionskassette meint chimäre DNA-Moleküle in denen eine für  
25 das dsRNA-Molekül (bzw. für einen der Stränge desselben) kodierende Nukleinsäuresequenz mit mindestens einem genetischen Kontrollelement (beispielsweise einem Promotor, Enhancer, Silencer, Splice-Donor oder -Akzeptor, Polyadenylierungssignal) derart verknüpft ist, das die Transkription des dsRNA-Moleküls (bzw. eines  
30 der Stränge desselben) in der eukaryotischen Zelle oder Organismus gewährleistet ist. Entsprechend vorteilhafte Konstruktionen sind weiter unten beschrieben. Eine Polyadenylierung ist möglich, jedoch nicht erforderlich, ebenso müssen keine Elemente zur Initiierung einer Translation vorhanden sein.

35 Soll das Expressionskonstrukt in eine Pflanze eingeführt und die dsRNA in plantae erzeugt werden, so sind pflanzenspezifische genetische Kontrollelemente (beispielsweise pflanzenspezifische Promotoren) bevorzugt. Die dsRNA kann jedoch auch in anderen Organismen oder in vitro erzeugt und dann in die Pflanze eingebracht werden.  
40

Unter einer funktionellen Verknüpfung versteht man zum Beispiel die sequentielle Anordnung eines Promotors mit der zu transkribierenden Nukleinsäuresequenz und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator und/oder Polyadenylierungssignalen derart, dass jedes der regulativen Elemente seine  
45

## 18

Funktion bei der Transkription der Nukleinsäuresequenz, je nach Anordnung der Nukleinsäuresequenzen erfüllen kann. Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die zu transkribierende Nukleinsäuresequenz hinter der als Promoter fungierenden Sequenz positioniert wird, so dass beide Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der transgen zu exprimierende Nukleinsäuresequenz geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare. In einer bevorzugten Ausführungsform wird die zu transkribierende Nukleinsäuresequenz so hinter dem Promotor lokalisiert, dass der Transkriptionsstart identisch ist mit dem gewünschten Beginn der dsRNA.

Die Herstellung einer funktionellen Verknüpfung als auch die Herstellung einer Expressionskassette kann mittels gängiger Rekombinations- und Klonierungstechniken realisiert werden, wie sie beispielsweise in Maniatis T, Fritsch EF und Sambrook J (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Silhavy TJ, Berman ML und Enquist LW (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Ausubel FM et al. (1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience und bei Gelvin et al. (1990) In: Plant Molecular Biology Manual beschrieben sind.

Unter einer Expressionskassette sind aber auch solche Konstruktionen zu verstehen, bei denen zum Beispiel eine Nukleinsäuresequenz kodierend für eines dsRNA derart hinter einen endogenen Promotor platziert werden, dass der gleiche Effekt auftritt. Beide Ansätze führen zu Expressionskassetten im Sinne der Erfindung.

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für das erfindungsgemäße Verfahren verwendet werden, solange sie die Expression in dem Zielorganismus gewährleisten. Darüberhinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

Es können weitere Promotoren funktionell mit der zu exprimierenden Nukleinsäuresequenz verknüpft sein, die eine Expression in weiteren Eukaryoten oder in Prokaryoten, wie zum Beispiel *E.coli* Bakterien ermöglichen.

5

Die in den erfindungsgemässen Expressionskassetten oder Vektoren enthaltenen Nukleinsäuresequenzen können mit weiteren genetischen Kontrollsequenzen neben einem Promoter funktionell verknüpft sein. Der Begriff der genetischen Kontrollsequenzen ist breit zu  
10 verstehen und meint all solche Sequenzen, die einen Einfluss auf das Zustandekommen oder die Funktion der erfindungsgemässen Expressionskassette haben. Genetische Kontrollsequenzen modifizieren zum Beispiel die Transkription in prokaryotischen oder eukaryotischen Organismen. Vorzugsweise umfassen die erfindungsgemässen Expressionskassetten 5'-stromaufwärts von der jeweiligen  
15 transgen zu exprimierenden Nukleinsäuresequenz einen pflanzenspezifischen Promoter und 3'-stromabwärts eine Terminatorsequenz als zusätzliche genetische Kontrollsequenz, sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils funktionell  
20 verknüpft mit der transgen zu exprimierenden Nukleinsäuresequenz.

Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Regionen, Introns oder nichtkodierende 3'-Region von Genen. Es ist gezeigt worden, dass diese eine signifikante Funktion  
25 nen bei der Regulation der Genexpression spielen können. Kontrollsequenzen umfassen ferner Polyadenylierungssignale sowie Terminatorsequenzen.

Die Expressionskassette kann vorteilhafterweise eine oder mehrere  
30 sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promoter enthalten, die eine erhöhte transgene Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der transgen zu exprimierenden Nukleinsäuresequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden, wie weitere regulatorische Elemente oder Terminatoren. Die transgen zu exprimierenden Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.  
35

Als Kontrollsequenzen sind weiterhin solche zu verstehen, die  
40 eine homologe Rekombination bzw. Insertion in das Genom eines Wirtsorganismus ermöglichen oder die Entfernung aus dem Genom erlauben. Methoden wie die cre/lox-Technologie erlauben eine gewebespezifische, unter Umständen induzierbare Entfernung der Expressionskassette aus dem Genom des Wirtsorganismus (Sauer B  
45 (1998) Methods. 14(4):381-92). Hier werden bestimmte flankierende

## 20

Sequenzen dem Zielgen angefügt (lox-Sequenzen), die später eine Entfernung mittels der cre-Rekombinase ermöglichen.

Bevorzugt kann die Expressionskassette, bestehend aus einer Ver-  
5 knüpfung von Promoter und zu transkribierender Nukleinsäurese-  
quenz, integriert in einem Vektor vorliegen und durch zum Bei-  
spiel Transformation - nach einem der unten beschriebenen Verfah-  
ren - in die eukaryotische Zelle oder Organismus eingebracht wer-  
den. Die nachfolgende Expression kann transient sein oder aber  
10 auch - bevorzugt - stabil nach Insertion (beispielsweise unter  
Verwendung von Selektionsmarkern) der Expressionskassetten in das  
Genom erfolgen. Bevorzugt wird das dsRNA-Expressionssystem stabil  
in das Genom - beispielsweise die chromosomale DNA oder die DNA  
der Organellen (z.B. der Plastiden, Mitochondrien etc.) - einer  
15 Zelle integriert.

Die Einführung einer erfindungsgemässen transgenen Expressions-  
kassette in einen Organismus oder Zellen, Geweben, Organe, Teile  
bzw. Samen desselben (bevorzugt in Pflanzen bzw. pflanzliche Zel-  
20 len, Gewebe, Organe, Teile oder Samen) kann vorteilhaft unter  
Verwendung von Vektoren realisiert werden, in denen die transge-  
nen Expressionskassetten enthalten sind. Vektoren können bei-  
spielhaft Plasmide, Cosmide, Phagen, Viren oder auch Agrobakte-  
rien sein. Die transgenen Expressionskassetten können in den Vek-  
25 tor (bevorzugt ein Plasmidvektor) über eine geeignete Restrikti-  
onsschnittstelle insertiert werden. Der entstandene Vektor wird  
zunächst in E.coli eingeführt. Korrekt transformierte E.coli wer-  
den selektioniert, gezüchtet und der rekombinante Vektor mit dem  
Fachmann geläufigen Methoden gewonnen. Restriktionsanalyse und  
30 Sequenzierung können dazu dienen, den Klonierungsschritt zu über-  
prüfen. Bevorzugt sind solche Vektoren, die eine stabile Integra-  
tion der Expressionskassette in das Wirtsgenom ermöglichen.

Die Herstellung eines transformierten Organismus (bzw. einer  
35 transformierten Zelle oder Gewebes) erfordert, dass die entspre-  
chende DNA (z.B. der Expressionsvektor) oder RNA in die entspre-  
chende Wirtszelle eingebracht wird. Für diesen Vorgang, der als  
Transformation (oder Transduktion bzw. Transfektion) bezeichnet  
wird, steht eine Vielzahl von Methoden zur Verfügung (Keown et  
40 al. (1990) Methods in Enzymology 185:527-537). So kann die DNA  
oder RNA beispielhaft direkt durch Mikroinjektion oder durch Bom-  
bardierung mit DNA-beschichteten Mikropartikeln eingeführt wer-  
den. Auch kann die Zelle chemisch, zum Beispiel mit Polyethylen-  
glycol, permeabilisiert werden, so dass die DNA durch Diffusion  
45 in die Zelle gelangen kann. Die DNA kann auch durch Protopla-  
stenfusion mit anderen DNA-enthaltenden Einheiten wie Minicells,  
Zellen, Lysosomen oder Liposomen erfolgen. Elektroporation ist

eine weitere geeignete Methode zur Einführung von DNA, bei der die Zellen reversibel durch einen elektrischen Impuls permeabilisiert werden. Entsprechende Verfahren sind beschrieben (beispielsweise bei Bilanz et al. (1991) *Gene* 100:247-250; Scheid et al. (1991) *Mol Gen Genet* 228:104-112; Guerche et al. (1987) *Plant Science* 52:111-116; Neuhaase et al. (1987) *Theor Appl Genet* 75:30-36; Klein et al. (1987) *Nature* 327:70-73; Howell et al. (1980) *Science* 208:1265; Horsch et al. (1985) *Science* 227:1229-1231; DeBlock et al. (1989) *Plant Physiology* 91:694-701; 10 *Methods for Plant Molecular Biology* (Weissbach and Weissbach, eds.) Academic Press Inc. (1988); and *Methods in Plant Molecular Biology* (Schuler and Zielinski, eds.) Academic Press Inc. (1989)).

- 15 Ein weiterer Gegenstand der Erfindung betrifft Zellen, die eines der erfindungsgemäßen dsRNA Moleküle, Expressionssysteme, Expressionskassetten oder Expressionsvektoren enthalten. Die Zelle kann von einem Organismus abgeleitet oder in diesem enthalten sein, meint aber auch einzellige Organismen wie Mikroorganismen. Die 20 Zelle kann prokaryotisch oder eukaryotischer Natur sein. Wobei das erfindungsgemäße Verfahren auf eukaryotische Organismen angewendet wird. Dennoch können prokaryotische Organismen die erfindungsgemäßen Expressionssysteme beispielsweise zum Zwecke der dsRNA-Produktion enthalten. Auch können prokaryotische Organismen, beispielsweise Agrobakterien, vorteilhaft als Vehikel für 25 die Transformation beispielsweise pflanzlicher Organismen eingesetzt werden.

Bevorzugte Prokaryoten sind vor allem Bakterien wie Bakterien der 30 Gattung *Escherichia*, *Corynebacterium*, *Bacillus*, *Clostridium*, *Proionibacterium*, *Butyrivibrio*, *Eubacterium*, *Lactobacillus*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Phaeodactylum*, *Colpidium*, *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodium* oder Cyanobakterien zum Beispiel der Gattung *Synechocystis*. Bevorzugt sind vor allem Mikroorganismen, welche zur Infektion von 35 Pflanzen und damit zur Übertragung der erfindungsgemäßen Konstrukte befähigt sind. Bevorzugte Mikroorganismen sind solche aus der Gattung *Agrobacterium* und insbesondere der Art *Agrobacterium tumefaciens*.

40

Eukaryotische Zellen und Organismen umfasst pflanzliche und tierische, nicht-menschliche Organismen und/oder Zellen sowie eukaryotische Mikroorganismen wie beispielsweise Hefen, Algen oder Pilze. Eine entsprechender transgener Organismus kann beispielsweise 45 weise durch Einführung der entsprechenden Expressionssysteme in

eine Zygote, Stammzelle, Protoplast oder eine andere geeignete von dem Organismus abgeleitete Zelle hergestellt werden.

- "Tierische Organismus" meint nicht-menschliche Vertebraten oder  
5 Invertebraten. Bevorzugte Vertebraten umfassen beispielsweise Fischearten, nicht-menschliche Säugetiere wie Rind, Pferd, Schaf, Ziege, Maus, Ratte oder Schwein, sowie Vögel wie Huhn oder Gans. Bevorzugte tierische Zellen umfassen CHO, COS, HEK293 Zellen. Invertebraten umfassen Nematoden oder andere Würmer sowie Insekten.  
10 Invertebraten umfassen Insektenzellen wie Drosophila S2 und Spodoptera Sf9 oder Sf21 Zellen.

- Bevorzugt sind ferner Nematoden, die in der Lage sind Tiere oder Menschen zu befallen, wie solche der Gattungen Ancylostoma, Ascaridia, Ascaris, Bunostomum, Caenorhabditis, Capillaria, Chabertia, Cooperia, Dictyocaulus, Haemonchus, Heterakis, Nematodirus, Oesophagostomum, Ostertagia, Oxyuris, Parascaris, Strongylus, Toxascaris, Trichuris, Trichostrongylus, Trichonema, Toxocara oder Uncinaria. Ferner bevorzugt sind solche, die in der Lage sind  
20 pflanzliche Organismen zu befallen wie beispielsweise Bursaphelenchus, Criconebella, Ditylenchus, Globodera, Helicotylenchus, Heterodera, Longidorus, Meloidogyne, Nacobbus, Paratylenchus, Pratylenchus, Radopholus, Rotelynchus, Tylenchus oder Xiphinema. Bevorzugte Insekten umfassen solche der Gattungen Coleoptera, Diptera, Lepidoptera und Homoptera.  
25

- Bevorzugte Pilze sind Aspergillus, Trichoderma, Ashbya, Neurospora, Fusarium, Beauveria oder weitere in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) auf Seite 15, Tabelle 6 beschriebene Pilze. Besonders bevorzugt ist der filamentöse Hemiascomycet Ashbya gossypii.  
30

- Bevorzugte Hefen sind Candida, Saccharomyces, Hansenula oder Pichia, besonders bevorzugt sind Saccharomyces cerevisiae oder Pichia pastoris (ATCC Accession No. 201178).  
35

- Als transgene Organismen bevorzugt sind vor allem pflanzliche Organismen. "Pflanzlicher Organismus" umfasst jeden Organismus, der zur Photosynthese befähigt ist, sowie die von diesem abgeleitete  
40 Zellen, Gewebe, Teile oder Vermehrungsgut (wie Samen oder Früchte). Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Einjährige, mehrjährige, monocotyledone und dicotyledone Pflanzen sowie Gymnospermen sind bevorzugt. Eingeschlossen sind  
45 reife Pflanze, Saatgut, Sprosse und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut (zum Beispiel Knollen, Samen oder Früchte) und Kulturen, zum Beispiel Zell- oder Kalluskulturen.

## 23

Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.

- 5 "Pflanze" im Rahmen der Erfindung meint alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Eingeschlossen unter dem Begriff sind die reifen Pflanzen, Saatgut, Sprossen und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut, Pflanzenorgane, Gewebe, Protoplasten, Kallus und andere
- 10 Kulturen, zum Beispiel Zellkulturen, sowie alle anderen Arten von Gruppierungen von Pflanzenzellen zu funktionellen oder strukturellen Einheiten. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.
- 15 "Pflanze" umfasst alle einjährigen und mehrjährige, monokotyledonen und dikotyledonen Pflanzen und schließt beispielhaft jedoch nicht einschränkend solche der Gattungen Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solarium, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Zea, Avena, Hordeum, Secale, Triticum, Sorghum, Picea und Populus ein.

- Bevorzugt sind Pflanzen nachfolgender Pflanzenfamilien: Amarantaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae,
- 30 Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Sterculiaceae, Tetragoniaceae, Theaceae, Umbelliferae.

- 35 Bevorzugte monokotyle Pflanzen sind insbesondere ausgewählt aus den monokotylen Kulturpflanzen, wie zum Beispiel der Familie der Gramineae wie Alfalfa, Reis, Mais, Weizen oder andere Getreidearten wie Gerste, Hirse, Roggen, Triticale oder Hafer sowie dem Zuckerrohr sowie alle Arten von Gräsern.

- 40 Die Erfindung wird ganz besonders bevorzugt aus dikotyledonen pflanzliche Organismen angewendet. Bevorzugte dikotyle Pflanzen sind insbesondere ausgewählt aus den dikotylen Kulturpflanzen, wie zum Beispiel



## 24

- Asteraceae wie Sonnenblume, Tagetes oder Calendula und andere mehr,
  - Compositae, besonders die Gattung Lactuca, ganz besonders die  
5 Art sativa (Salat) und andere mehr,
  - Cruciferae, besonders die Gattung Brassica, ganz besonders die Arten napus (Raps), campestris (Rübe), oleracea cv Tastie (Kohl), oleracea cv Snowball Y (Blumenkohl) und oleracea cv Emperor (Broccoli) und weitere Kohlarten; und der Gattung Arabidopsis, ganz besonders die Art thaliana sowie Kresse oder Canola und andere mehr,  
10
  - Cucurbitaceae wie Melone, Kürbis oder Zucchini und andere mehr,  
15
  - Leguminosae besonders die Gattung Glycine, ganz besonders die Art max (Sojabohne) Soja sowie Alfalfa, Erbse, Bohnengewächsen oder Erdnuss und andere mehr
  - 20 - Rubiaceae, bevorzugt der Unterklasse Lamiidae wie beispielsweise Coffea arabica oder Coffea liberica (Kaffestrauch) und andere mehr,
  - Solanaceae besonders die Gattung Lycopersicon, ganz besonders  
25 die Art esculentum (Tomate) und die Gattung Solanum, ganz besonders die Art tuberosum (Kartoffel) und melongena (Aubergine) sowie Tabak oder Paprika und andere mehr,
  - Sterculiaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Theobroma cacao (Kakaostrauch) und andere mehr,  
30
  - Theaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Camellia sinensis oder Thea sinensis (Teestrauch) und andere mehr,  
35
  - Umbelliferae, besonders die Gattung Daucus (ganz besonders die Art carota (Karrotte)) und Apium (ganz besonders die Art graveolens dulce (Selderie)) und andere mehr; und die Gattung Capsicum, ganz besonders die Art annum (Pfeffer) und andere mehr,  
40
- sowie Lein, Soya, Baumwolle, Hanf, Flachs, Gurke, Spinat, Möhre, Zuckerrübe und den verschiedenen Baum-, Nuss- und Weinarten, insbesondere Banane und Kiwi.
- 45 Umfasst sind ferner Schmuckpflanzen, Nutz- oder Zierbäume, Blumen, Schnittblumen, Sträucher oder Rasen. Beispielhaft aber nicht einschränkend seien zu nennen Angiospermen, Bryophyten wie zum

## 25

Beispiel Hepaticae (Leberblümchen) und Musci (Moose); Pteridophyten wie Farne, Schachtelhalm und Lycopoden; Gymnospermen wie Koniferen, Cycaden, Ginkgo und Gnetalen, die Familien der Rosaceae wie Rose, Ericaceae wie Rhododendrons und Azaleen, Euphorbiaceae wie Weihnachtssterne und Kroton, Caryophyllaceae wie Nelken, Solanaceae wie Petunien, Gesneriaceae wie das Usambaraveilchen, Balsaminaceae wie das Springkraut, Orchidaceae wie Orchideen, Iridaceae wie Gladiolen, Iris, Freesie und Krokus, Compositae wie Ringelblume, Geraniaceae wie Geranien, Liliaceae wie der Drachenbaum, Moraceae wie Ficus, Araceae wie Philodendron und andere mehr.

Pflanzliche Organismen im Sinne der Erfindung sind weiterhin weitere photosynthetisch aktive befähigte Organismen, wie zum Beispiel Algen, Cyanobakterien sowie Moose. Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung Haematococcus, Phaedactylum tricornatum, Volvox oder Dunaliella. Insbesondere bevorzugt ist Synechocystis.

20 Am meisten bevorzugt sind

- a) Pflanzen, die zur Ölproduktion geeignet sind, wie beispielsweise Raps, Sonnenblume, Sesam, Färberdistel (*Carthamus tinctorius*), Ölbaum, Soja, Mais, Erdnuß, Rizinus, Ölpalme, Weizen, Kakaotrauch oder verschiedene Nussarten wie beispielsweise Walnuss, Kokosnuß oder Mandel. Unter diesen wieder besonders bevorzugt sind dikotyledonen Pflanzen, insbesondere Raps, Soja und Sonnenblume.
- 25 b) Pflanzen, die der Stärkeproduktion dienen, wie beispielsweise Mais, Weizen oder Kartoffel.
- c) Pflanzen, die als Nahrungs- und/oder Futtermittel und/oder Nutzpflanze genutzt werden und bei denen eine Resistenz gg. Pathogene vorteilhaft wäre, wie beispielsweise Gerste, Roggen, Reis, Kartoffel, Baumwolle, Flachs, Lein.
- 35 d) Pflanzen, die zur Produktion von Feinchemikalien wie beispielsweise Vitaminen und/oder Carotinoiden dienen können, wie beispielsweise Raps.
- 40

Die in den Verfahren verwendeten Organismen werden je nach Wirtsorganismus in dem Fachmann bekannter Weise angezogen bzw. gezüchtet. Mikroorganismen werden in der Regel in einem flüssigen Medium, das eine Kohlenstoffquelle meist in Form von Zuckern, eine Stickstoffquelle meist in Form von organischen Stickstoffquellen wie Hefeextrakt oder Salzen wie Ammoniumsulfat, Spurenelemente

wie Eisen-, Mangan-, Magnesiumsalze und gegebenenfalls Vitamine enthält, bei Temperaturen zwischen 0°C und 100°C, bevorzugt zwischen 10°C bis 60°C unter Sauerstoffbegasung angezogen. Dabei kann der pH der Nährflüssigkeit auf einen festen Wert gehalten werden, 5 das heißt während der Anzucht reguliert werden oder nicht. Die Anzucht kann batch weise, semi batch weise oder kontinuierlich erfolgen. Nährstoffe können zu Beginn der Fermentation vorgelegt oder semikontinuierlich oder kontinuierlich nach gefüttert werden.

10

Nachfolgende Anwendung des erfindungsgemäßen Verfahrens seien beispielhaft jedoch nicht einschränkend zu nennen:

#### I. Pflanzenbiotechnologie

15

Bevorzugt wird das erfindungsgemäße Verfahren im Rahmen der Pflanzenbiotechnologie zur Erzeugung von Pflanzen mit vorteilhaften Eigenschaften eingesetzt. So kann Eignung der Pflanzen oder deren Samen als Nahrungs- oder Futtermittel verbessert werden, 20 beispielsweise über eine Veränderung der Zusammensetzungen und/oder des Gehalt an Metaboliten, insbesondere Proteinen, Ölen, Vitaminen und/oder Stärke. Auch können Wachstumsrate, Ertrag oder die Resistenz gegen biotische oder abiotische Stressfaktoren erhöht werden. Nachfolgende Anwendungen im Bereich der Pflanzenbiotechnologie sind insbesondere vorteilhaft. Die angegebenen möglichen Zielgene sind beispielhaft jedoch nicht einschränkend zu verstehen: 25

1. Verbesserter Schutz gegen abiotische Stressfaktoren (Hitze, 30 Kälte, Trockenheit, erhöhte Feuchtigkeit, Umweltgifte, UV-Strahlung). Bevorzugt werden Gene in ihrer Expression vermindert, die an Stressreaktionen beteiligt sind.

2. Modifikation der Zusammensetzung und/oder des Gehaltes an 35 Fettsäuren, Lipiden oder Ölen

Eine Veränderung des Fettsäuregehaltes oder der Fettsäurezusammensetzung, vorzugsweise in einer Ölfrucht wie Raps oder Sonnenblume, kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen der Fettsäurebiosynthese vorzugsweise ausgewählt aus der Gruppe bestehend aus Genen kodierend für Acetyltransacylasen, Acyltransportproteinen ("acyl carrier protein"), Desaturasen wie Stearyl-desaturasen oder mikrosomale  $\Delta 12$ -Desaturasen insbesondere Fad2-1 Gene, 40 Malonyltransacylase,  $\beta$ -Ketoacyl-ACP-synthetasen, 3-Keto-ACP-reduktasen, Enoyl-ACP-hydrasen, Thioesterasen wie Acyl-ACP-thioesterases, Enoyl-ACP-reduktasen. Verschiedene weitere vor- 45

teilhafte Ansätze zur Modifizierung der Lipidzusammensetzung sind beschrieben (Shure M et al. (1983) Cell 35:225-233; Preiss et al. (1987) Tailoring Genes for Crop Improvement (Bruening et al., eds.), Plenum Press, S.133-152; Gupta et al. (1988) Plant Mol Biol. 10:215-224; Olive et al. (1989) Plant Mol Biol 12:525-538; Bhattacharyya et al. (1990) Cell 60:155-122; Dunwell JM (2000) J Exp Botany 51Spec No:487-96; Brar DS et al. (1996) Biotech Genet Eng Rev 13:167-79; Kishore GM und Somerville CR (1993) Curr Opin Biotech 4(2):152-8). Bevorzugt sind insbesondere Fad2 Gene (z.B. beschrieben durch Genbank Acc.-Nr.: AF124360 (*Brassica carinata*), AF042841 (*Brassica rapa*), L26296 (*Arabidopsis thaliana*), A65102 (*Corylus avellana*)). Weitere vorteilhafte Gene und Verfahren zur Modifikation des Lipidgehaltes sind beispielsweise beschrieben in US 5,530,192 und WO 94/18337. Ein erhöhter Lipidgehalt kann auch erreicht werden durch Verminderung des Stärkegehaltes beispielsweise infolge verminderter Expression von von Enzymen des Kohlenhydratstoffwechsels (z.B. ADP-Glucosepyrophosphorylasen).

### 20 3. Modifikation der Kohlenhydratzusammensetzung

Eine Modifikation der Kohlenhydratzusammensetzung kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen des Kohlenhydratstoffwechsels oder der Kohlenhydratbiosynthese, beispielsweise der Biosynthese von Amylose, Pektinen, Cellulose oder Zellwandkohlenhydraten. Dadurch kann eine Vielzahl zellulärer Prozesse (Reifung, Halfestigkeit, Stärkezusammensetzung oder -gehalt etc.) in vorteilhafter Weise beeinflusst werden. Als Zielgene seien beispielhaft jedoch nicht einschränkend zu nennen Phosphorylasen, Stärkesynthetasen, Q-Enzyme, Sucrose-6-phosphatsynthetasen, Sucrose-6-phosphatphosphatasen, ADP-Glucosepyrophosphorylasen, Branching-Enzyme, Debranching-Enzyme sowie diverse Amylasen. Die entsprechenden Gene sind beschrieben (Dunwell JM (2000) J Exp Botany 51Spec No:487-96; Brar DS et al. (1996) Biotech Genet Eng Rev 13:167-79; Kishore GM und Somerville CR (1993) Curr Opin Biotech 4(2):152-8). Vorteilhafte Gene zur beeinflussung des Kohlenhydratstoffwechsels - insbesondere der Stärkebiosynthese - sind beschrieben in WO 92/11375, WO 92/11376, US 5, 365,016 und WO 95/07355.

### 4. Veränderung der Farbe oder Pigmentierung

Veränderung der Farbe oder Pigmentierung vorzugsweise von Zierpflanzen kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen der Flavonoid-Biosynt-

hese wie beispielsweise Chalconsynthasen, Chalconisomerasen, Phenylalaninammonialyasen, Dehydrokaempferol(flavone)hydroxylasen wie Flavanon-3-hydroxylasen oder Flavanon-2-hydroxylasen, Dihydroflavonolreduktasen, Dihydroflavanol-2-hydroxylasen, Flavonoid-3'-hydroxylasen, Flavonoid-5'-hydroxylasen, Flavonoidglycosyltransferasen (z.B. Glucosyltransferasen wie UDPG:Flavonoid- 3-O-glucosyltransferasen, UDPG:Flavonol-7-O-glucosyltransferasen oder Rhamnosyltransferasen), Flavonoidmethyltransferasen (wie z.B. SAM:Anthocyanidin-3-(p-coumaroyl)-rutinosid-5-glucosid-3',5'-O-methyltransferasen) und Flavonoidacyltransferasen (Hahlbrock (1981) Biochemistry of Plants, Vol.7, Conn (Ed.); Weiring and de Vlamming (1984) "Petunia", KC Sink (Ed.), Springer-Verlag, New York). Geeignet sind insbesondere die in EP-A1 522 880 beschriebenen Sequenzen.

#### 5. Verminderung des Gehaltes von Speicherproteinen

Die Verminderung der Genexpression von Genen kodierend für Speicherproteine (infolge SP) hat zahlreiche Vorteile, wie beispielsweise Verminderung des allergenen Potentials oder Veränderung in der Zusammensetzung oder Menge anderer Metabolite. Speicherproteine sind u.a. beschrieben in EP-A 0 591 530, WO 87/47731, WO 98/26064, EP-A 0 620 281; Kohno-Murase J et al. (1994) Plant Mol Biol 26(4): 1115-1124.

SP dienen zur Speicherung von Kohlenstoff, Stickstoff und Schwefel, die für das schnelle heterotrophe Wachstum bei Keimung von Samen oder Pollen benötigt werden. Sie haben meist keine enzymatische Aktivität. SP werden dabei nur im Embryo während der Samenentwicklung synthetisiert und akkumulieren dabei zum einen in Proteinspeichervakuolen (PSV) von unterschiedlich differenzierten Zellen im Embryo bzw. Endosperm.

"Speicherprotein" meint allgemein ein Protein, das mindestens eine der nachfolgenden wesentlichen Eigenschaften aufweist:

i) Speicherproteine werden im wesentlichen nur im Embryo während der Samenentwicklung exprimiert. "Im wesentlichen" bedeutet dabei, dass in dem besagten Stadium mindestens 50%, bevorzugt mindestens 70%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% der Gesamtexpression über die Lebensdauer einer Pflanze hinweg stattfindet.

## 29

ii) Speicherproteine werden während der Keimung des Samens wieder abgebaut. Dabei beträgt der Abbau während der Keimung mindestens 20%, bevorzugt mindestens 50%, ganz besonders bevorzugt mindestens 80%.

5

iii) Speicherproteine machen einen wesentlichen Anteil am Gesamtproteingehalt des nicht-keimenden Samens aus. Bevorzugt macht das Speicherprotein in dem nicht-keimenden Samen der Wildtyp-Pflanze mehr als 5 Gew.% des Gesamtproteins aus, besonders bevorzugt mindestens 10 Gew.%, ganz besonders bevorzugt mindestens 20 Gew.%, am meisten bevorzugt mindestens 30 Gew.-%.

10

Bevorzugt weisen Speicherproteine 2 oder alle der oben genannten wesentlichen Eigenschaften i), ii) oder iii) auf.

15

Speicherproteine können in Untergruppen entsprechend weiterer charakteristischer Eigenschaften, wie beispielsweise ihrem Sedimentationskoeffizienten oder der Löslichkeit in verschiedenen Lösungen (Wasser, Salzlösung, Alkohol) aufgeteilt werden. Die Bestimmung des Sedimentationskoeffizienten kann in der dem Fachmann vertrauten Weise mittels Ultrazentrifugation durchgeführt werden (z.B. beschrieben bei Correia JJ (2000) Methods in Enzymology 321:81-100).

20

25

Insgesamt können vier grosse Genfamilien für Speicherproteine aufgrund ihrer Sequenzen zugeordnet werden: 2S-Albumine (Napin-ähnlich), 7S-Globuline (Phaseolin-ähnlich), 11S/12S-Globuline (Legumin-/Cruciferin-ähnlich) und die Zein-Prolamine.

30

2S Albumine sind weit verbreitet in Samen von Dikotyledonen, einschliesslich wichtiger kommerzieller Pflanzenfamilien wie Fabaceae (z.B. Sojabohne), Brassicaceae (z.B. Raps), Euphorbiaceae (z.B. Rizinus) oder Asteraceae (z.B. Sonnenblume). 2S Albumine sind kompakte globuläre Proteine mit konservierten Cysteinresten, die oft Heterodimere bilden.

35

7S-Globuline liegen in trimerer Form vor und enthalten keine Cysteinreste. Nach ihrer Synthese werden sie wie die 2S-Albumine in kleinere Fragmente gespalten und glykosyliert. Trotz Unterschiede in der Polypeptidgrösse sind die verschiedenen 7S-Globuline hoch konserviert und gehen vermutlich wie die 2S-Albumine auf ein gemeinsames Vorläuferprotein zurück. Die 7S-Globuline sind nur in geringen Mengen in Monokotyledonen vorhanden. In Dikotyledonen ist ihr Anteil immer kleiner verglichen mit den 11S/12S-Globulinen.

40

45

## 30

11S/12S-Globuline stellen neben den 2S-Albuminen die Hauptfraktion der Speicherproteine in Dikotyledonen. Die hohe Ähnlichkeit der verschiedenen 11S-Globuline aus verschiedenen Pflanzengattungen lassen wiederum auf einen gemeinsamen Vorläuferprotein in der Evolution schliessen.

Bevorzugt ist das Speicherprotein ausgewählt aus den Klassen der 2S-Albumine (Napin-ähnlich), 7S-Globuline (Phaseolin-ähnlich), 11S/12S-Globuline (Legumin-/Cruciferin-ähnlich) oder Zein-Prolamine.

Besonders bevorzugte 2S-Albumine umfassen

i) 2S-Albumine aus Arabidopsis, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 2, 4, 6 oder 8, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 1, 3, 5 oder 7 kodierten Proteine,

ii) 2S-Albumine aus Arten der Gattung Brassica, wie beispielsweise Brassica napus, Brassica nigra, Brassica juncea, Brassica oleracea oder Sinapis alba, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 32, 34, 36, 38, 40, 46 oder 48, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 31, 33, 35, 37, 39, 45 oder 47 kodierten Proteine,

iii) 2S-Albumine aus Soja, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 42 oder 44, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 41 oder 43 kodierten Proteine,

iv) 2S-Albumine aus Sonnenblume (Helianthus annuus), ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 50 oder 52, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 49 oder 51 kodierten Proteine,

sowie die entsprechenden Homologen und funktionellen Äquivalente zu i) oder ii) oder iii) oder iv) aus identischen oder anderen Pflanzenarten, insbesondere Raps, Sonnenblume, Lein, Sesam, Färberdistel, Ölbaum, Soja oder verschiedene Nussarten. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 2S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Wasser aus.

## 31

- 5 Funktionelle Äquivalente der 2S-Albumine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer  
10 der Proteinsequenzen mit der SEQ ID NO: 2, 4, 6, 8, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 oder 52 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge  
15 der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 2S-Speicherproteins auf.
- 15 Besonders bevorzugte 7S-Globuline umfassen solche aus Arabidopsis oder Soja, ganz besonders bevorzugt die Proteine mit der SEQ ID NO: 94 oder 96, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 93 oder 95 kodierten Proteine.  
20 Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 7S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Salzlösung aus. Als weitere charakteristische Eigenschaft können 7S-Globuline keine Cysteinreste enthalten.
- 25 Funktionelle Äquivalente der 7S-Globuline haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer  
30 der Proteinsequenzen mit der SEQ ID NO: 94 oder 96 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die  
35 gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 7S-Speicherproteins auf.
- 40 Besonders bevorzugte 11S/12S-Globuline umfassen bevorzugt 11S-Globuline aus Raps, Soja und Arabidopsis insbesondere
- 45 i) 11S-Globuline aus Raps mit der SEQ ID NO: 10, 12, 14, 16 oder 18, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 9, 11, 13, 15 oder 17 kodierten Proteine,



## 32

ii) die 11S-Globuline aus Soja mit der SEQ ID NO: 20, 22, 24, 26 oder 28, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 19, 21, 23, 25 oder 27 kodierten Proteine,

5

iii) die 11S-Globuline aus *Arabidopsis thaliana* mit der SEQ ID NO: 60, 62, 64, 66, 68 oder 70 am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 59, 61, 63, 65, 67 oder 69 kodierten Proteine,

10

sowie die entsprechenden Homologen und funktionellen Äquivalente aus anderen Pflanzenarten, insbesondere Raps, Sonnenblume, Lein, Sesam, Färberdistel, Ölbaum, Soja oder verschiedene Nussarten, wie beispielsweise das Sonnenblume 11S Speicherprotein (SEQ ID NO: 30), insbesondere das durch die Nukleinsäuresequenz gemäß SEQ ID NO: 29 kodierte Protein. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 11S- oder 12S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Salzlösung (PBS; phosphatgepufferte Salzlösung) und/oder eine schlechte Löslichkeit in Wasser aus.

15

20

25

30

35

Funktionelle Äquivalente der 11S- oder 12S Albumine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer der Proteinsequenzen mit der SEQ ID NO: 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 60, 62, 64, 66, 68 oder 70 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 11S- oder 12S-Speicherproteins auf.

40

45

Besonders bevorzugte Zein-Prolamine umfassen bevorzugt solche aus monokotyledonen Pflanzen, insbesondere Mais, Reis, Hafer, Gerste oder Weizen. Ganz besonders bevorzugt sind die Mais Zein-Prolamine beschrieben durch SEQ ID NO: 98, 100, 102 oder 104 - insbesondere die durch SEQ ID NO 97, 99, 101 oder 103 kodierten Protein -, das Reis Prolamin gemäß SEQ ID NO: 106 - insbesondere das durch SEQ ID NO 105 kodierte Protein -, das Hafer Prolamin gemäß SEQ ID NO: 108 - insbesondere das durch SEQ ID NO 107 kodierte Proteine-, das Gerste Prolamin gemäß SEQ ID NO: 110 und/oder 111 - insbesondere das durch SEQ ID

## 33

NO 109 kodierte Protein - und das das Weizen Prolamin gemäß  
SEQ ID NO: 113 - insbesondere das durch SEQ ID NO 112 ko-  
dierte Protein. Funktionelle Äquivalente zeichnen sich bevor-  
zugt durch eine Löslichkeit in 70%iger ethanolischer Lösung  
5 und eine schlechte Löslichkeit in Wasser oder Salzlösung aus.

Funktionelle Äquivalente der Zein-Prolamine haben in einer  
weiteren bevorzugten Ausführungsform eine Homologie von min-  
destens 60%, bevorzugt mindestens 80%, ganz besonders bevor-  
zugt mindestens 90%, am meisten bevorzugt mindestens 95% zu  
10 einer der Proteinsequenzen mit der SEQ ID NO: 98, 100, 102,  
104, 106, 108, 110, 111 oder 113 wobei die Homologie sich be-  
vorzugt über eine Länge von mindestens 30 Aminosäuren, bevor-  
zugt mindestens 50 Aminosäuren besonders bevorzugt über 100  
15 Aminosäuren, am meisten bevorzugt über die gesamte Länge der  
jeweiligen Proteine erstreckt, und weisen die gleichen we-  
sentlichen Eigenschaften eines Speicherproteins und - bevor-  
zugt- die charakteristischen Eigenschaften eines Zein-Prola-  
mine auf.

20 Funktionelle Äquivalente meint insbesondere natürliche oder  
künstliche Mutationen der obengenannten Speicherproteine so-  
wie homologe Polypeptide aus anderen Pflanzen, die die glei-  
chen wesentlichen und - bevorzugt - charakteristischen Eigen-  
25 schaften aufweisen. Bevorzugt sind homologe Polypeptide aus  
oben beschriebenen bevorzugten Pflanzen. Die zu den im Rahmen  
dieser Erfindung offenbarten Speicherproteinen homologen Se-  
quenzen aus anderen Pflanzen - beispielsweise solchen deren  
genomische Sequenz ganz oder teilweise bekannt ist, wie bei-  
30 spielsweise aus *Arabidopsis thaliana*, *Brassica napus*, *Nico-*  
*tiana tabacum* oder *Solanum tuberosum* - durch Homologiever-  
gleiche aus Datenbanken auffinden. können z.B. durch Daten-  
banksuche oder Durchmustern von Gen-Banken - unter Verwendung  
der beispielhaft aufgeführten Speicherprotein-Sequenzen als  
35 Suchsequenz bzw. Sonde - leicht aufgefunden werden.

Mutationen umfassen Substitutionen, Additionen, Deletionen,  
Inversion oder Insertionen eines oder mehrerer Aminosäurere-  
ste.

40 Ein weiterer Gegenstand der Erfindung umfasst ein zumindest  
teilweise doppelsträngiges Ribonukleinsäuremolekül, wobei das  
doppelsträngige Ribonukleinsäuremolekül umfasst

45 i) einen "sense"-RNA-Strang umfassend mindestens zwei Ribo-  
nukleotidsequenzabschnitte, wobei jeweils mindestens ei-  
ner dieser Ribonukleotidsequenzabschnitte im wesentlichen

## 34

identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 59, 61, 63, 65, 67, 69, 71, 93, 95, 97, 99, 101, 103, 105, 107, 109 oder 112 oder eines funktionellen Äquivalentes derselben, wobei jedoch nicht alle Ribonukleotidsequenzabschnitte zu dem "sense"-RNA-Transkript eines einzigen einer Speicherprotein-Nukleinsäuresequenz identisch sind, und

ii) einen "antisense"-RNA-Strang, der zu dem RNA-sense-Strang unter i) im wesentlichen komplementären ist.

Bevorzugt haben zumindest zwei der Speicherprotein-Nukleinsäuresequenzen, zu deren "sense"-RNA-Transkript die besagten Ribonukleotidsequenzabschnitte im wesentlichen identisch sind, untereinander eine Homologie von unter 90%, bevorzugt unter 80%, ganz besonders bevorzugt unter 60% am meisten bevorzugt unter 50% über die gesamte Länge ihrer kodierenden Nukleotidsequenz.

In einer weiteren bevorzugten Ausführungsform enthält die dsRNA mehrere Sequenzabschnitte, die eine gleichzeitige Suppression mehrerer Speicherproteine, bevorzugt von Speicherproteinen aus verschiedenen Klassen - wie beispielsweise einem 2S-Albumin, 7S-Globuline, 11S/12S-Globulin oder die Zein-Prolamine - bewirken.

Am meisten bevorzugt sind doppelsträngige RNA Moleküle beschrieben durch die Ribonukleinsäuresequenz gemäß SEQ ID NO: 84, 86 oder 88. Diese werden bevorzugt kodiert durch Nukleotidsequenzen entsprechend SEQ ID NO: 83, 85 oder 87.

5. Erreichen einer Resistenz gegen pflanzliche Pathogene

Eine Resistenz gegen pflanzliche Pathogene wie Arachniden, Pilze, Insekten, Nematoden, Protozoen, Viren, Bakterien und Krankheiten kann erreicht werden durch Verminderung der Genexpression von Genen, die für das Wachstum, Überleben, bestimmte Entwicklungsstufen (beispielsweise Verpuppung) oder die Vermehrung eines bestimmten Pathogens essentiell sind. Eine entsprechende Verminderung kann eine vollständige Inhibition vorgenannter Schritte aber auch eine Verzögerung derselben bewirken. Dies können pflanzliche Gene sein, die dem Pathogen beispielsweise das Eindringen ermöglichen, können aber auch pathogen-eigene Gene sein. Bevorzugt ist die dsRNA

## 35

gg. Gene des Pathogens gerichtet. Als anti-Pathogenes Agens kann dabei die dsRNA selber, jedoch auch die Expressionssysteme, Expressionskassetten oder transgenen Organismen wirken. Pflanzen können beispielsweise mit geeigneten Formulierungen vorgenannter Agentien behandelt, beispielsweise be-  
5 sprüht oder estäubt werden. Die Pflanzen selber können jedoch in Form eines transgenen Organismus die Agentien beinhalten und diese - beispielsweise in Form eines Fraßgiftes - an die Pathogene weitergeben. Verschiedene essentielle Gene diverser  
10 Pathogene sind dem Fachmann bekannt (z.B. für Nematodenresistenz WO 93/10251, WO 94/17194).

Am meisten bevorzugt als Pathogen sind Pilzpathogene wie *Phytophthora infestans*, *Fusarium nivale*, *Fusarium graminearum*,  
15 *Fusarium culmorum*, *Fusarium oxysporum*, *Blumeria graminis*, *Magnaporthe grisea*, *Sclerotinia sclerotium*, *Septoria nodorum*, *Septoria tritici*, *Alternaria brassicae*, *Phoma lingam*, bakterielle Pathogene wie *Corynebacterium sepedonicum*, *Erwinia carotovora*, *Erwinia amylovora*, *Streptomyces scabies*, *Pseudomonas syringae* pv. *tabaci*, *Pseudomonas syringae* pv. *phaseolicola*, *Pseudomonas syringae* pv. *tomato*, *Xanthomonas campestris* pv. *malvacearum* und *Xanthomonas campestris* pv. *oryzae*, und Nematoden wie *Globodera rostochiensis*, *G. pallida*, *Heterodera schachtii*, *Heterodera avenae*, *Ditylenchus dipsaci*, *Anguina tritici* und *Meloidogyne hapla*.  
20  
25

Eine Virusresistenz kann beispielsweise durch Verminderung der Expression eines viralen Hüllproteins, einer viralen Replikase, einer viralen Protease etc. erreicht werden. Zahlreiche Pflanzenviren und entsprechende Zielgene sind dem  
30 Fachmann bekannt.

#### 6. Verhinderung von Halmbruch

35 Eine verminderte Anfälligkeit gegen Halmbruch kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen des Kohlenhydratstoffwechsels (s.o.). Vorteilhafte Gene sind beschrieben (u.a. WO 97/13865) und umfassen gewebespezifische Polygalacturonasen oder Cellulasen.  
40

#### 7. Verzögerung der Fruchtreifung

Eine verzögerte Fruchtreifung kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen ausgewählt aus der Gruppe bestehend aus Polygalacturonasen, Pectinesterasen,  $\beta$ -(1-4)glucanasen (Cellulasen),  $\beta$ -Galactanasen ( $\beta$ -Galactosidasen), oder Gene der Ethylenbiosynthese wie  
45

## 36

- 1-Aminocyclopropan-1-carboxylatsynthase, Gene der Carotinoidbiosynthese wie z.B. Gene der Prephytoen- oder Phytoenbiosynthese beispielsweise Phytoendesaturasen. Weitere vorteilhafte Gene sind beispielsweise in WO 91/16440, WO 91/05865, WO 5 91/16426, WO 92/17596, WO 93/07275 oder WO 92/04456.
8. Erzielen einer männlichen Sterilität ("male sterility"). Entsprechende Zielgene sind beschrieben u.a. in WO 94/29465, WO89/10396, WO 92/18625.
- 10 9. Verminderung unerwünschter oder toxischer Pflanzeninhaltsstoffe wie z.B. Glucosinolaten. Entsprechende Zielgene sind beschrieben (u.a. in WO 97/16559).
- 15 10. Verzögerung von Alterserscheinungen. Entsprechende Zielgene sind u.a. Cinnamoyl-CoA:NADPH-Reduktasen oder Cinnamoylalkoholdehydrogenasen. Weitere Zielgene sind beschrieben (u.a. in WO 95/07993).
- 20 11. Modifikation der Lignifikation und/oder des Ligningehaltes vor allem in Baumarten. Entsprechende Zielgene sind beschrieben u.a. in WO 93/05159, WO 93/05160.
- 25 12. Modifikation des Faseranteils in Nahrungsmitteln vorzugsweise in Samen durch Verminderung der Expression der Coffeinsäure-O-methyltransferase oder der Cinnamoylalkoholdehydrogenase.
13. Modifikation der Faserqualität in Baumwolle. Entsprechende Zielgene sind beschrieben u.a. in US 5,597,718.
- 30 14. Verminderung der Stoßanfälligkeit von beispielsweise Kartoffeln durch Verminderung beispielsweise der Polyphenoloxidase (WO 94/03607) etc.
- 35 15. Steigerung der Vitamin E Biosynthese beispielsweise durch Verminderung der Expression von Genen aus dem Homogentisatabauweg wie z.B. der Homogentisat-1,2-dioxygenase (HGD; EC-Nr.: 1.13.11.5), der Maleylacetoacetatisomerase (MAAI; EC-Nr.: 5.2.1.2.) oder der Fumarylacetoacetathydrolase (FAAH; 40 EC-Nr.: 3.7.1.2).

Ein weiterer Gegenstand der Erfindung umfasst ein zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

- i) einen "sense"-RNA-Strang umfassend mindestens zwei Ribonukleotidsequenzabschnitte, wobei jeweils mindestens einer dieser Ribonukleotidsequenzabschnitte im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines Gens aus dem Homogentisatabbauweg gemäß SEQ ID NO: 115, 116, 118 oder 120 oder eines funktionellen Äquivalentes derselben, wobei jedoch nicht alle Ribonukleotidsequenzabschnitte zu dem "sense"-RNA-Transkript eines einzigen einer Speicherprotein-Nukleinsäuresequenz identisch sind, und
- ii) einen "antisense"-RNA-Strang, der zu dem RNA-sense-Strang unter i) im wesentlichen komplementären ist.
16. Verminderung des Nikotingehaltes beispielsweise in Tabak durch verminderte Expression beispielsweise der N-Methylputrescinoxidase und der Putrescin-N-methyltransferase.
17. Verminderung des Coffeingehaltes in der Kaffeebohne (*Coffea arabica*) durch Verminderung der Genexpression von Genen der Coffeinbiosynthese wie 7-Methylxanthine-3-methyltransferase.
18. Verminderung des Theophyllin-Gehaltes im Tee (*Camellia sinensis*) durch Verminderung der Genexpression von Genen der Theophyllin-Biosynthese wie beispielsweise 1-Methylxanthin-3-methyltransferase.
19. Erhöhung des Methioningehaltes durch Verminderung der Threoninbiosynthese, beispielsweise durch Verminderung der Expression der Threoninsynthase (Zeh M et al. (2001) Plant Physiol 127(3):792-802).

Weitere Beispiele für vorteilhafte Gene sind zum Beispiel genannt bei Dunwell JM, Transgenic approaches to crop improvement, J Exp Bot. 2000;51 Spec No; Seite 487-96.

Jede der oben genannten Anwendungen kann als solche isoliert angewendet werden. Natürlich können auch mehr als eine der oben genannten Ansätze gleichzeitig angewendet werden. Dabei wird bei allen Anwendungen die Expression von mindestens zwei unterschiedlichen Zielgenen, wie oben definiert, vermindert. Diese Zielgene können dabei aus einer einzigen für eine Anwendung bevorzugten Gruppe von Genen stammen oder aber auch unterschiedlichen Anwendungsgruppen zugeordnet sein.

Zur Anwendung der erfindungsgemäßen Verfahren stehen dem Fachmann geläufige Werkzeuge, wie Expressionsvektoren mit für Pflanzen geeigneten Promotoren, sowie Verfahren zur Transformation und Regeneration von Pflanzen zur Verfügung. Pflanzenspezifische

- 5 Promotoren meint grundsätzlich jeden Promotor, der die Expression von Genen, insbesondere Fremdgenen, in Pflanzen oder Pflanzenteilen, -zellen, -geweben, -kulturen steuern kann. Dabei kann die Expression beispielsweise konstitutiv, induzierbar oder entwicklungsabhängig sein. Bevorzugt sind:

10

a) Konstitutive Promotoren

- "Konstitutive" Promotoren meint solche Promotoren, die eine Expression in zahlreichen, bevorzugt allen, Geweben über einen  
 15 größeren Zeitraum der Pflanzenentwicklung, bevorzugt zu allen Zeitpunkten der Pflanzenentwicklung, gewährleisten (Benfey et al. (1989) EMBO J 8:2195-2202). Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Insbesondere bevorzugt ist der  
 20 Promotor des 35S-Transkriptes des CaMV Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228) oder der 19S CaMV Promotor (US 5,352,605; WO 84/02913; Benfey et al. (1989)  
 25 EMBO J 8:2195-2202). Ein weiterer geeigneter konstitutiver Promotor ist der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028), der LeguminB-Promotor (GenBank Acc.-Nr. X03677), der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), den Ubiquitin 1 Promotor (Christensen et al. (1992) Plant Mol Biol 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), den SmaS Promotor, den Cinnamylal-  
 30 koholdehydrogenase-Promotor (US 5,683,439), die Promotoren der vakuolärer ATPase Untereinheiten oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991), sowie weitere Promotoren von Genen, deren konstitutive Expression in Pflanzen dem Fachmann bekannt ist.

40 b) Gewebespezifische Promotoren

Bevorzugt sind ferner Promotoren mit Spezifitäten für die Ant-  
 heren, Ovarien, Blüten, Blätter, Stengel, Wurzeln und Samen.

- 45 Samenspezifische Promotoren wie zum Beispiel der Promotor des Phaseolins (US 5,504,200; Bustos MM et al. (1989) Plant Cell 1(9):839-53), des 2S Albumingens (Joseffson LG et al. (1987) J

## 39

- Biol Chem 262:12196-12201), des Legumins (Shirsat A et al. (1989) Mol Gen Genet 215(2): 326-331), des USP (unknown seed protein; Bäumlein H et al. (1991) Mol Gen Genet 225(3):459-67), des Napin Gens (US 5,608,152; Stalberg K et al. (1996) L Planta 199:515-519), des Saccharosebindeproteins (WO 00/26388) oder der Legumin B4-Promotor (LeB4; Bäumlein H et al. (1991) Mol Gen Genet 225: 121-128; Bäumlein et al. (1992) Plant Journal 2(2):233-9; Fiedler U et al. (1995) Biotechnology (NY) 13(10):1090f), der Oleosin-Promoter aus Arabidopsis (WO 98/45461), der Bce4-Promoter aus Brassica (WO 91/13980). Weitere geeignete samenspezifische Promotoren sind die der Gene kodierend für das "High Molecular Weight Glutenin" (HMWG), Gliadin, Verzweigungsenzym, ADP Glucose Pyrophosphatase (AG-Pase) oder die Stärkesynthase. Bevorzugt sind ferner Promotoren, die eine samenspezifische Expression in Monokotyledonen wie Mais, Gerste, Weizen, Roggen, Reis etc. erlauben. Vorteilhaft eingesetzt werden können der Promoter des lpt2 oder lpt1-Gen (WO 95/15389, WO 95/23230) oder die Promotoren beschrieben in WO 99/16890 (Promotoren des Hordein-Gens, des Glutelin-Gens, des Oryzin-Gens, des Prolamin-Gens, des Gliadin-Gens, des Glutelin-Gens, des Zein-Gens, des Kasirin-Gens oder des Secalin-Gens). Weitere samenspezifische Promotoren sind beschrieben in WO89/03887.
- 25 Knollen-, Speicherwurzel- oder Wurzel-spezifische Promotoren wie beispielsweise der Patatin Promotor Klasse I (B33), der Promotor des Cathepsin D Inhibitors aus Kartoffel.
- 30 Blattspezifische Promotoren wie Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubisco (Ribulose-1,5-bisphosphatcarboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al. (1989) EMBO J 8:2445-2451).
- 35 Blüten-spezifische Promotoren wie beispielsweise der Phytoen Synthase Promotor (WO 92/16635) oder der Promotor des P-rr Gens (WO 98/22593).
- 40 - Antheren-spezifische Promotoren wie den 5126-Promotor (US 5,689,049, US 5,689,051), den glob-1 Promotor und den  $\gamma$ -Zein Promotor.

## c) Chemisch induzierbare Promotoren

- 45 Die Expressionskassetten können auch einen chemisch induzierbaren Promotor enthalten (Übersichtsartikel: Gatz et al. (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108), durch den die



Expression des exogenen Gens in der Pflanze zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige Promotoren, wie z.B. der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366), durch Salicylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzolsulfonamid-induzierbarer Promotor (EP 0 388 186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure induzierbarer Promotor (EP 0 335 528) bzw. ein durch Ethanol- oder Cyclohexanon-induzierbarer Promotor (WO 93/21334) können ebenfalls verwendet werden.

d) Stress- oder Pathogen-induzierbare Promotoren

Ferner sind Promotoren bevorzugt, die durch biotischen oder abiotischen Stress induziert werden wie beispielsweise der pathogen-induzierbare Promotor des PRP1-Gens (Ward et al. (1993) Plant Mol Biol 22:361-366), der hitzeinduzierbare hsp70- oder hsp80-Promoter aus Tomate (US 5,187,267), der kälteinduzierbare alpha-Amylase Promoter aus der Kartoffel (WO 96/12814), der licht-induzierbare PPDK Promotor oder der verwundungsinduzierte pinII-Promoter (EP375091).

Pathogen-induzierbare Promotoren umfassen die von Genen, die infolge eines Pathogenbefalls induziert werden wie beispielsweise Gene von PR-Proteinen, SAR-Proteinen,  $\beta$ -1,3-Glucanase, Chitinase usw. (beispielsweise Redolfi et al. (1983) Neth J Plant Pathol 89:245-254; Uknes, et al. (1992) The Plant Cell 4:645-656; Van Loon (1985) Plant Mol Viral 4:111-116; Marineau et al. (1987) Plant Mol Biol 9:335-342; Matton et al. (1987) Molecular Plant-Microbe Interactions 2:325-342; Somssich et al. (1986) Proc Natl Acad Sci USA 83:2427-2430; Somssich et al. (1988) Mol Gen Genetics 2:93-98; Chen et al. (1996) Plant J 10:955-966; Zhang and Sing (1994) Proc Natl Acad Sci USA 91:2507-2511; Warner, et al. (1993) Plant J 3:191-201; Siebertz et al. (1989) Plant Cell 1:961-968(1989).

Umfasst sind auch verwundungs-induzierbare Promotoren wie der des pinII Gens (Ryan (1990) Ann Rev Phytopath 28:425-449; Duan et al. (1996) Nat Biotech 14:494-498), des wun1 und wun2-Gens (US 5,428,148), des win1- und win2-Gens (Stanford et al. (1989) Mol Gen Genet 215:200-208), des Systemin (McGurl et al. (1992) Science 225:1570-1573), des WIP1-Gens (Rohmeier et al. (1993) Plant Mol Biol 22:783-792; Eckelkamp et al. (1993) FEBS Letters 323:73-76), des MPI-Gens (Corderok et al. (1994) The Plant J 6(2):141-150) und dergleichen.

## 41

## e) Entwicklungsabhängige Promotoren

- Weitere geeignete Promotoren sind beispielsweise fruchtreifungsspezifische Promotoren, wie beispielsweise der fruchtreifungsspezifische Promotor aus Tomate (WO 94/21794, EP 409 625). Entwicklungsabhängige Promotoren schließt zum Teil die Gewebespezifischen Promotoren ein, da die Ausbildung einzelner Gewebe naturgemäß entwicklungsabhängig erfolgt.

Besonders bevorzugt sind konstitutive sowie samenspezifische Promotoren.

- Genetische Kontrollsequenzen umfassen auch weitere Promotoren, Promotorelemente oder Minimalpromotoren, die die expressionsteuernden Eigenschaften modifizieren können. So kann durch genetische Kontrollsequenzen zum Beispiel die gewebespezifische Expression zusätzlich abhängig von bestimmten Stressfaktoren erfolgen. Entsprechende Elemente sind zum Beispiel für Wasserstress, Abscisinsäure (Lam E und Chua NH, J Biol Chem 1991; 266(26): 17131-17135) und Hitzestress (Schoffl F et al., Molecular & General Genetics 217(2-3):246-53, 1989) beschrieben.

- Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Regionen, Introns oder nichtkodierende 3'-Region von Genen wie beispielsweise das Actin-1 Intron, oder die Adh1-S Introns 1, 2 und 6 (allgemein: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994)). Es ist gezeigt worden, dass diese eine signifikante Funktionen bei der Regulation der Genexpression spielen können. So wurde gezeigt, dass 5'-untranslatierte Sequenzen die transiente Expression heterologer Gene verstärken können. Beispielhaft für Translationsverstärker sei die 5'-Leadersequenz aus dem Tabak-Mosaik-Virus zu nennen (Gallie et al. (1987) Nucl Acids Res 15:8693-8711) und dergleichen. Sie können ferner die Gewebsspezifität fördern (Rouster J et al. (1998) Plant J 15:435-440).

- Als Kontrollsequenzen geeignete Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA Polyadenylierungssignale aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACHS entsprechen (Gielen et al. (1984) EMBO J 3:835 ff) oder funktionelle Äquivalente davon. Beispiele für besonders geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopalin-Synthase)-Terminator.

- Eine Expressionskassetten und die von ihr abgeleiteten Vektoren können weitere Funktionselemente enthalten. Der Begriff Funktionselement ist breit zu verstehen und meint all solche Elemente,

die einen Einfluss auf Herstellung, Vermehrung oder Funktion der erfindungsgemässen Expressionskassetten, Vektoren oder transgenen Organismen haben. Beispielhaft aber nicht einschränkend seien zu nennen:

5

a) Selektionsmarker, die eine Resistenz gegen einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456), Antibiotika oder Biozide, bevorzugt Herbizide, wie zum Beispiel Kanamycin, G 418, Bleomycin, Hygromycin, oder Phosphino-

10 notricin etc. verleihen. Besonders bevorzugte Selektionsmarker sind solche die eine Resistenz gegen Herbizide verleihen. Beispielhaft seien genannt: DNA Sequenzen, die für Phosphinothricinacetyltransferasen (PAT) kodieren und Glutaminsynthaseinhibitoren inaktivieren (bar und pat Gen), 5-Enolpyru-

15 vylshikimat-3-phosphatsynthasegene (EPSP Synthasegene), die eine Resistenz gegen Glyphosat® (N-(phosphonomethyl)glycin) verleihen, das für das Glyphosat® degradierende Enzyme kodierende gox Gen (Glyphosatoxidoreduktase), das deh Gen (kodierend für eine Dehalogenase, die Dalapon inaktiviert), Sulfonylurea- und Imidazolinon inaktivierende Acetolactatsynthasen

20 sowie bxn Gene, die für Bromoxynil degradierende Nitrilaseenzyme kodieren, das aasa-Gen, das eine Resistenz gegen das Antibiotikum Apectinomycin verleiht, das Streptomycinphosphotransferase (SPT) Gen, das eine Resistenz gegen Streptomycin

25 gewährt, das Neomycinphosphotransferase (NPTII) Gen, das eine Resistenz gegen Kanamycin oder Geneticidin verleiht, das Hygromycinphosphotransferase (HPT) Gen, das eine Resistenz gegen Hygromycin vermittelt, das Acetolactatsynthase Gen (ALS), das eine Resistenz gegen Sulfonylharnstoff-Herbizide verleiht

30 (z.B. mutierte ALS-Varianten mit z.B. der S4 und/oder Hra Mutation).

b) Reportergene, die für leicht quantifizierbare Proteine kodieren und über Eigenfarbe oder Enzymaktivität eine Bewertung

35 der Transformationseffizienz oder des Expressionsortes oder -zeitpunktes gewährleisten. Ganz besonders bevorzugt sind dabei Reporter-Proteine (Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) wie das "green fluorescence protein" (GFP) (Sheen et al. (1995) Plant Journal 8(5):777-784),

40 die Chloramphenicoltransferase, eine Luziferase (Ow et al. (1986) Science 234:856-859), das Aequorin-Gen (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), die  $\beta$ -Galactosidase, ganz besonders bevorzugt ist die  $\beta$ -Glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).

## 43

- c) Replikationsursprünge, die eine Vermehrung der erfindungsge-  
mässen Expressionskassetten oder Vektoren in zum Beispiel  
E.coli gewährleisten. Beispielhaft seien genannt ORI (origin  
of DNA replication), der pBR322 ori oder der P15A ori (Sam-  
brook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed.  
5 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,  
1989).
- d) Elemente, die für eine Agrobakterium vermittelte Pflanzen-  
10 transformation erforderlich sind, wie zum Beispiel die rechte  
oder linke Begrenzung der T-DNA oder die vir-Region.

Zur Selektion erfolgreich homolog rekombinierter oder auch trans-  
formierter Zellen ist es in der Regel erforderlich, einen selek-  
15 tionierbaren Marker zusätzlich einzuführen, der den erfolgreich  
rekombinierten Zellen eine Resistenz gegen ein Biozid (zum Bei-  
spiel ein Herbizid), einen Metabolismusinhibitor wie 2-Desoxyglu-  
cose-6-phosphat (WO 98/45456) oder ein Antibiotikum verleiht. Der  
Selektionsmarker erlaubt die Selektion der transformierten Zellen  
20 von untransformierten (McCormick et al. (1986) Plant Cell Reports  
5:81-84).

Verschiedene Methoden und Vektoren zum Einschleusen von Genen in  
das Genom von Pflanzen sowie zur Regeneration von Pflanzen aus  
25 Pflanzengeweben oder Pflanzenzellen sind bekannt (Plant Molecular  
Biology and Biotechnology (CRC Press, Boca Raton, Florida), Kapi-  
tel 6/7, S. 71-119 (1993); White FF (1993) Vectors for Gene  
Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Enginee-  
ring and Utilization, Hrsgb.: Kung und Wu R, Academic Press,  
30 15-38; Jenes B et al. (1993) Techniques for Gene Transfer, in:  
Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.:  
Kung und R. Wu, Academic Press, S.128-143; Potrykus (1991) Annu  
Rev Plant Physiol Plant Molec Biol 42:205-225; Halford NG, Shewry  
PR (2000) Br Med Bull 56(1):62-73). Dazu zählen beispielhaft die  
35 oben erwähnten. Bei Pflanzen werden dabei die beschriebenen Me-  
thoden zur Transformation und Regeneration von Pflanzen aus  
Pflanzengeweben oder Pflanzenzellen zur transienten oder stabilen  
Transformation genutzt. Geeignete Methoden sind vor allem die  
Protoplastentransformation durch Polyethylenglykol-induzierte  
40 DNA-Aufnahme, die Liposomen vermittelte Transformation (wie z.B.  
in US 4,536,475 beschrieben), biolistische Verfahren mit der Gen-  
kanone ("particle bombardment" Methode; Fromm ME et al. (1990)  
Bio/Technology. 8(9):833-9; Gordon-Kamm et al. (1990) The Plant  
Cell 2:603), die Elektroporation, die Inkubation trockener Em-  
45 bryonen in DNA-haltiger Lösung und die Mikroinjektion. Im Falle  
dieser "direkten" Transformationsmethoden sind keine besonderen  
Anforderungen an das verwendete Plasmid gestellt. Einfache Plas-

mide wie die der pUC-Reihe, pBR322, M13mp Reihe, pACYC184 etc. können verwendet werden. Sollen vollständige Pflanzen aus den transformierten Zellen regeneriert werden, so ist es erforderlich, das sich auf dem Plasmid ein zusätzliches selektionierbares 5 Markergen befindet.

Neben diesen "direkten" Transformationstechniken kann eine Transformation auch durch bakterielle Infektion mittels *Agrobacterium* (z.B. EP 0 116 718), virale Infektion mittels viraler Vektoren 10 (EP 0 067 553; US 4,407,956; WO 95/34668; WO 93/03161) oder mittels Pollen (EP 0 270 356; WO 85/01856; US 4,684,611) durchgeführt werden.

Die für die *Agrobacterium*-Transformation meist verwendeten 15 Stämme *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* eine auch durch bakterielle Infektion mittels enthalten ein Plasmid (Ti bzw. Ri Plasmid), das auf die Pflanze nach *Agrobacterium*-Infektion übertragen wird. Ein Teil dieses Plasmids, genannt T-DNA (transferred DNA), wird in das Genom der Pflanzenzelle inte- 20 griert. Alternativ können durch *Agrobacterium* auch binäre Vektoren (Mini-Ti-Plasmide) auf Pflanzen übertragen und in deren Genom integriert werden. Die *Agrobacterium*-vermittelte Transformation ist am besten für dicotyledone, diploide Pflanzenzellen geeignet, wohingegen die direkten Transformationstechniken sich für jeden 25 Zelltyp eignen. Verfahren zur *Agrobacterium* vermittelten Transformation sind beispielsweise beschrieben bei Horsch RB et al. (1985) Science 225:1229f. Werden *Agrobacterien* verwendet, so ist die Expressionskassette in spezielle Plasmide zu integrieren, entweder in einen Zwischenvektor (englisch: shuttle or interme- 30 diate vector) oder einen binären Vektor. Wird ein Ti oder Ri Plasmid zur Transformation verwendet werden soll, ist zumindest die rechte Begrenzung, meistens jedoch die rechte und die linke Begrenzung der Ti oder Ri Plasmid T-DNA als flankierende Region mit der einzuführenden Expressionskassette verbunden.

35 Für die *Agrobacterium* Transformation werden bevorzugt binäre Vektoren verwendet. Binäre Vektoren können sowohl in *E.coli* als auch in *Agrobacterium* replizieren. Sie enthalten in der Regel ein Selektionsmarkergen und einen Linker oder Polylinker flankiert von 40 der rechten und linken T-DNA Begrenzungssequenz. Sie können direkt in *Agrobacterium* transformiert werden (Holsters et al. (1978) Mol Gen Genet 163:181-187). Das Selektionsmarkergen erlaubt eine Selektion transformierter *Agrobacteria* und ist zum Beispiel das nptII Gen, das eine Resistenz gegen Kanamycin ver- 45 leiht. Das in diesem Fall als Wirtsorganismus fungierende *Agrobacterium* sollte bereits ein Plasmid mit der vir-Region enthalten. Diese ist für die Übertragung der T-DNA auf die pflanzliche

## 45

Zelle erforderlich. Ein so transformiertes Agrobacterium kann zur Transformation pflanzlicher Zellen verwendet werden. Die Verwendung von T-DNA zur Transformation pflanzlicher Zellen ist intensiv untersucht und beschrieben (EP 120 516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblas-  
5 nary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblas-  
serdam, Chapter V; An et al. (1985) EMBO J 4:277-287). Verschiedene binäre Vektoren sind bekannt und teilweise kommerziell erhältlich wie zum Beispiel pBI101.2 oder pBIN19 (Clontech Laboratories, Inc. USA; Bevan et al. (1984) Nucl Acids Res 12:8711),  
10 pBinAR, pPZP200 oder pPTV.

Die mit einem solchen Vektor transformierten Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen, wie z.B. von Raps, verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in  
15 einer Agrobakterienlösung gebadet und anschliessend in geeigneten Medien kultiviert werden. Die Transformation von Pflanzen durch Agrobakterien ist beschrieben (White FF, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu,  
20 Academic Press, 1993, S. 15 - 38; Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, S.128-143; Potrykus (1991) Annu Rev Plant Physiol Plant  
25 Molec Biol 42:205- 225). Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die integriert die oben beschriebenen erfindungsgemässen Expressionssysteme enthalten.

30 Stabil transformierte Zellen d.h. solche, die die eingeführte DNA integriert in die DNA der Wirtszelle enthalten, können von untransformierten selektioniert werden, wenn ein selektionierbarer Marker Bestandteil der eingeführten DNA ist. Als Marker kann beispielhaft jedes Gen fungieren, dass eine Resistenz gegen Antibio-  
35 tika oder Herbizide (wie Kanamycin, G418, Bleomycin, Hygromycin oder Phosphinotricin etc.) zu verleihen vermag (s.o.). Transformierte Zellen, die ein solches Markergen exprimieren, sind in der Lage, in der Gegenwart von Konzentrationen eines entsprechenden Antibiotikums oder Herbizides zu überleben, die einen untransfor-  
40 mierten Wildtyp abtöten. Beispiel sind oben genannt und umfassen bevorzugt das bar Gen, dass Resistenz gegen das Herbizid Phosphinotricin verleiht. (Rathore KS et al. (1993) Plant Mol Biol 21(5):871-884), das nptII Gen, dass Resistenz gegen Kanamycin verleiht, das hpt Gen, das Resistenz gegen Hygromycin verleiht,  
45 oder das EPSP-Gen, das Resistenz gegen das Herbizid Glyphosat verleiht. Der Selektionsmarker erlaubt die Selektion von transformierten Zellen von untransformierten (McCormick et al. (1986)

Plant Cell Reports 5:81-84). Die erhaltenen Pflanzen können in üblicher Weise gezüchtet und gekreuzt werden. Zwei oder mehr Generationen sollten vorzugsweise kultiviert werden, um sicherzustellen, dass die genomische Integration stabil und vererblich ist.

Sobald eine transformierte Pflanzenzelle hergestellt wurde, kann eine vollständige Pflanze unter Verwendung von dem Fachmann bekannten Verfahren erhalten werden. Hierbei geht man beispielhaft von Kalluskulturen aus. Aus diesen noch undifferenzierten Zellmassen kann die Bildung von Spross und Wurzel in bekannter Weise induziert werden. Die erhaltenen Sprösslinge können ausgepflanzt und gezüchtet werden. Entsprechende Verfahren sind beschrieben (Fennell et al. (1992) Plant Cell Rep. 11: 567-570; Stoeger et al (1995) Plant Cell Rep. 14:273-278; Jahne et al. (1994) Theor Appl Genet 89:525-533).

Die Wirksamkeit der Expression der transgen exprimierten Nukleinsäuren kann beispielsweise *in vitro* durch Sprossmeristemvermehrung unter Verwendung einer der oben beschriebenen Selektionsmethoden ermittelt werden. Zudem kann eine in Art und Höhe veränderte Expression eines Zielgens und die Auswirkung auf den Phänotyp der Pflanze an Testpflanzen in Gewächshausversuchen getestet werden.

25

## II. Medizinische Anwendungen

Die erfindungsgemäß bereitgestellten dsRNA, Expressionssysteme oder Organismen eignen sich zur Herstellung von Arzneimitteln zur Behandlung von menschlichen und tierischen Erkrankungen. Für eine effizient Therapie ist es oft unzureichend nur ein einzelnes Zielgen zu vermindern. Das erfindungsgemäße Verfahren eignet sich insbesondere zur Behandlung von

35 - Pathogenbefall, wie beispielsweise virale oder bakterielle Erkrankungen. In diesen Fällen führen Ansätze, die lediglich gegen ein molekulares Ziel gerichtet sind, oft zu der Ausbildung von Resistenzen. Eine Kombinationstherapie, die mehrere Ziele abdeckt, ist jedoch kompliziert zu koordinieren und v.a. nur sehr aufwendig in klinischen Experimenten zu evaluieren. Das erfindungsgemäße Verfahren ermöglicht hier eine vorteilhafte Alternative. Die inhibitorische dsRNA kann dabei in der dem Fachmann geläufigen Weise appliziert werden. dsRNA verfügt über eine

45 erstaunliche Stabilität und effiziente Wirkung und kann beispielsweise durch Verfütterung entsprechender dsRNA exprimierenden Bakterien appliziert werden. Das Verfahren

## 47

5 eignet sich insbesondere zur Behandlung von viralen Infektionen z.B. mit dem "human immunodeficiency virus" (HIV), indem gleichzeitig die Expression von mindestens zwei viralen Gene vermindert wird, beispielsweise bei HIV von Genen wie gp41, die für den Zelleintritt verantwortlich sind, und der viralen Replikase oder reversen Transkriptase.

- 10 - Behandlung von Krebs (beispielsweise solider Tumore und/oder Leukämien). Zahlreiche potentialle Zielgene sind hier dem Fachmann bekannt (z.B. Oncogene wie ABL1, BCL1, BCL2, BCL6, CBFA2, CBL, CSF1R, ERBA, ERBB, EBRB2, FGR, FOS, FYN, HRAS, JUN, LCK, LYN, MYB, MYC, NRAS, RET oder SRC; Tumorsuppressorgene wie BRCA1 oder BRCA2; Adhäsionsmoleküle; Cyclinekinasen und deren Inhibitoren).
- 15

20 Weitere potentiell mit dem erfindungsgemäßen Verfahren behandelbare Erkrankungen und die entsprechenden Zielgene sind dem Fachmann ohne weiteres zugänglich und umfassen beispielsweise Erkrankungen des Herz/Kreislaufsystems wie Bluthochdruck, Erkrankungen des zentralen oder peripheren Nervensystems wie Alzheimer, Parkinson oder multiple Sklerose usw. Auch ist es durch das erfindungsgemäße Verfahren möglich, mehr als eine Erkrankung parallel zu behandeln, wie beispielsweise ein

25 Herzkreislauferkrankung und eine Erkrankung des zentralen Nervensystems, was durch klassische Ansätze nicht möglich ist. Derartige Ansätze sind v.a. bei multiplen Erkrankungen wie sie oft im fortgeschrittenen Alter auftreten vorteilhaft. Beispielfhaft sei die parallele Behandlung von Bluthochdruck und z.B. Alzheimer oder seniler Demenz zu nennen. Dabei kann

30 dieser Anwendungen als solche isoliert angewendet werden. Natürlich können auch mehr als eine der oben genannten Ansätze gleichzeitig angewendet werden. Dabei wird bei allen Anwendungen die Expression von mindestens zwei unterschiedlichen Zielgenen vermindert. Diese Zielgene können dabei aus der für

35 eine Anwendung bevorzugten Gruppe von Genen stammen oder aber auch unterschiedlichen Anwendungsgruppen zugeordnet sein.

### III. Biotechnologische Anwendungen

40

Das erfindungsgemäße Verfahren läßt sich vorteilhaft in biotechnologischen Verfahren anwenden. Beispielfhaft jedoch nicht einschränkend sei zu nennen die Optimierung von Stoffwechselwegen in fermentativ genutzten Hefen, Pilzen oder anderen eukaryotischen Mikroorganismen oder Zellen zur Herstellung von

45 Feinchemikalien wie Aminosäuren (z.B. Lysin oder Methionin), Vitaminen (wie Vitamin B2, Vitamin C, Vitamin E), Carotinoiden



den, Ölen und Fetten, polyungesättigten Fettsäuren, Biotin usw. Dabei kann dieser Anwendungen als solche isoliert angewendet werden. Natürlich können auch mehr als eine der oben genannten Ansätze gleichzeitig angewendet werden. Dabei wird bei allen Anwendungen die Expression von mindestens zwei unterschiedlichen Zielgenen vermindert. Diese Zielgene können dabei aus der für eine Anwendung bevorzugten Gruppe von Genen stammen oder aber auch unterschiedlichen Anwendungsgruppen zugeordnet sein.

10

Als Vektoren zur Expression in E.coli sind bevorzugt pQE70, pQE60 und pQE-9 (QIAGEN, Inc.); pBluescript Vektoren, Phagescript Vektoren, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene Cloning Systems, Inc.); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia Biotech, Inc.).

Bevorzugte Vektoren zur eukaryotischen Expression umfassen pWLNE0, pSV2CAT, pOG44, pXT1 und pSG (Stratagene Inc.); pSVK3, pBPV, pMSG und pSVL (Pharmacia Biotech, Inc.). Als induzierbare Vektoren seien pTet-tTak, pTet-Splice, pcDNA4/TO, pcDNA4/TO / LacZ, pcDNA6/TR, pcDNA4/TO/Myc-His /LacZ, pcDNA4/TO/Myc-His A, pcDNA4/TO/Myc-His B, pcDNA4/TO/Myc-His C, pVgRXR (Invitrogen, Inc.) oder die pMAM-Serie (Clontech, Inc.; GenBank Accession No.: U02443) zunennen. Diese stellen bereits das induzierbare regulatorische Kontrollelement beispielsweise für eine chemisch, induzierbare Expression eines DSBI-Enzyms zur Verfügung. In diese Vektoren kann die Nukleinsäuresequenz kodierend für ein DSBI-Enzym direkt insertiert werden.

30 Vektoren für die Expression in Hefe umfassen beispielhaft pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, PHIL-D2, PHIL-S1, pPIC3SK, pPIC9K, und PA0815 (Invitrogen, Inc.).

35 Vorteilhafte Kontrollsequenzen sind beispielsweise die gram-positiven Promotoren amy und SPO2, und die Hefe- oder Pilzpromotoren ADC1, MFa, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH.

Klonierungsvektoren und Techniken zur genetischen Manipulation von Ciliaten und Algen sind dem Fachmann bekannt (WO 98/01572; Falciatore et al. (1999) Marine Biotechnology 1(3):239-251; Dunahay et al. (1995) J Phycol 31:10004-1012).

Als Selektionsmarker können prinzipiell viele der auch für Pflanzen bevorzugten Selektionssysteme verwendet werden. Insbesondere bevorzugt sind für Säugerzelle die Neomycin (G418) Resistenz, die Hygromycin-Resistenz, die Zeocin-Resistenz oder die Puromycin-Re-

49

sistenz. Für Prokaryoten ist insbesondere die Ampicillin-Resistenz, die Kanamycin-Resistenz oder die Tetracyclin-resistent bevorzugt.

5 Prinzipiell sind für die Transformation tierischer Zelle oder von Hefezellen ähnliche Verfahren wie für die "direkte" Transformation von pflanzlichen Zellen anzuwenden. Insbesondere Verfahren wie die Calciumphosphat oder Liposomen vermittelte Transformation oder aber Elektroporation sind bevorzugt.

10

Ein weiterer Gegenstand der Erfindung betrifft die Verwendung der erfindungsgemässen, transgenen Organismen und der von ihnen abgeleitete Zellen, Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln, Blätter etc.-, und trans-

15 genes Vermehrungsgut wie Saaten oder Früchte, zur Herstellung von Nahrungs- oder Futtermitteln, Pharmazeutika oder Feinchemikalien, wie beispielsweise Enzymen, Vitaminen, Aminosäuren, Zucker, Fettsäuren, natürliche und synthetische Geschmacks-, Aroma- und Farbstoffe. Besonders bevorzugt ist die Produktion von Triacylglyceriden, Lipiden, Ölen, Fettsäuren, Stärke, Tocopherolen und Tocotrienolen sowie Carotinoiden. Von Menschen und Tieren verzehrbare erfindungsgemässe, genetisch veränderte Pflanzen können auch beispielsweise direkt oder nach an sich bekannter Aufbereitung als Nahrungsmittel oder Futtermittel verwendet werden.

25

30

35

40

45

## Sequenzen

1. SEQ ID NO: 1  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 1 (GenBank Acc.-No.: M22032)
2. SEQ ID NO: 2  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 1
3. SEQ ID NO: 3  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 3 (GenBank Acc.-No.: M22035)
4. SEQ ID NO: 4  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 3
5. SEQ ID NO: 5  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 2 (GenBank Acc.-No.: M22034)
6. SEQ ID NO: 6  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 2
7. SEQ ID NO: 7  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 4 (GenBank Acc.-No.: M22033)
8. SEQ ID NO: 8  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 4
9. SEQ ID NO: 9  
Nukleinsäuresequenz kodierend für B.napus Cruciferin Spei-  
cherprotein (GenBank Acc.-No.: X59294)
10. SEQ ID NO: 10  
Proteinsequenz kodierend für B.napus Cruciferin Speicherpro-  
tein
11. SEQ ID NO: 11  
Nukleinsäuresequenz kodierend für Brassica napus Cruciferin  
(GenBank Acc.-No.: X14555)
12. SEQ ID NO: 12  
Proteinsequenz kodierend für Brassica napus Cruciferin

## 51

13. SEQ ID NO: 13  
Nukleinsäuresequenz kodierend für B.napus BnC2 Cruciferin  
Speicherprotein (GenBank Acc.-No.: X59295)
- 5 14. SEQ ID NO: 14  
Proteinsequenz kodierend für B.napus BnC2 Cruciferin Spei-  
cherprotein
15. SEQ ID NO: 15  
10 partielle Nukleinsäuresequenz kodierend für B.napus Crucife-  
rin cru4 subunit (GenBank Acc.-No.: X57848)
16. SEQ ID NO: 16  
partielle Proteinsequenz kodierend für B.napus Cruciferin  
15 cru4 subunit
17. SEQ ID NO: 17  
Nukleinsäuresequenz kodierend für B.napus cru1 Cruciferin  
subunit (GenBank Acc.-No.: X62120)  
20
18. SEQ ID NO: 18  
Proteinsequenz kodierend für B.napus cru1 Cruciferin subunit
19. SEQ ID NO: 19  
25 Nukleinsäuresequenz kodierend für Glycinin A-1a-B-x subunit  
aus des Sojabohne (GenBank Acc.-No.: M36686)
20. SEQ ID NO: 20  
Proteinsequenz kodierend für Glycinin A-1a-B-x subunit aus  
30 des Sojabohne
21. SEQ ID NO: 21  
Nukleinsäuresequenz kodierend für Sojabohne Glycinin subunit  
G2 (GenBank Acc.-No.: X15122)  
35
22. SEQ ID NO: 22  
Proteinsequenz kodierend für Sojabohne Glycinin subunit G2
23. SEQ ID NO: 23  
40 Nukleinsäuresequenz kodierend für Sojabohne A5A4B3 Glycinin  
subunits (GenBank Acc.-No.: X02626)
24. SEQ ID NO: 24  
Proteinsequenz kodierend für Sojabohne A5A4B3 Glycinin sub-  
45 units

## 52

25. SEQ ID NO: 25  
Nukleinsäuresequenz kodierend für Sojabohne (G.max) Glycinin  
Speicherprotein subunit A3-B4 (GenBank Acc.-No.: M10962)
- 5 26. SEQ ID NO: 26  
Proteinsequenz kodierend für Sojabohne (G.max) Glycinin Spei-  
cherprotein subunit A3-B4
27. SEQ ID NO: 27  
10 Nukleinsäuresequenz kodierend für Sojabohne Glycinin subunit  
G3 (GenBank Acc.-No.: X15123)
28. SEQ ID NO: 28  
Proteinsequenz kodierend für Sojabohne Glycinin subunit G3  
15
29. SEQ ID NO: 29  
Nukleinsäuresequenz kodierend für Sonnenblume 11S Speicher-  
protein (G3-D1) (GenBank Acc.-No.: M28832)
- 20 30. SEQ ID NO: 30  
Proteinsequenz kodierend für Sonnenblume 11S Speicherprotein  
(G3-D1)
31. SEQ ID NO: 31  
25 Nukleinsäuresequenz kodierend für Raps (B.napus) Napin (Gen-  
Bank Acc.-No.: J02586)
32. SEQ ID NO: 32  
Proteinsequenz kodierend für Raps (B.napus) Napin  
30
33. SEQ ID NO: 33  
Nukleinsäuresequenz kodierend für Brassica juncea 2S Spei-  
cherprotein (GenBank Acc.-No.: X65040)
- 35 34. SEQ ID NO: 34  
Proteinsequenz kodierend für Brassica juncea 2S Speicherpro-  
tein
35. SEQ ID NO: 35  
40 Nukleinsäuresequenz kodierend für Brassica oleracea 2S Spei-  
cherprotein (GenBank Acc.-No.: X65038)
36. SEQ ID NO: 36  
Proteinsequenz kodierend für Brassica oleracea 2S Speicher-  
45 protein

## 53

37. SEQ ID NO: 37  
Nukleinsäuresequenz kodierend für Brassica napus cv. Topas Napin (GenBank Acc.-No.: U04945)
- 5 38. SEQ ID NO: 38  
Proteinsequenz kodierend für Brassica napus cv. Topas Napin
39. SEQ ID NO: 39  
partielle Nukleinsäuresequenz kodierend für Sinapis alba sin1  
10 Speicherprotein (GenBank Acc.-No.: X91799)
40. SEQ ID NO: 40  
partielle Proteinsequenz kodierend für Sinapis alba sin1  
Speicherprotein  
15
41. SEQ ID NO: 41  
Nukleinsäuresequenz kodierend für Sojabohne (Glycine max) napin-type 2S Albumin 1 (GenBank Acc.-No.: U71194)
- 20 42. SEQ ID NO: 42  
Proteinsequenz kodierend für Sojabohne (Glycine max) napin-type 2S Albumin 1
43. SEQ ID NO: 43  
25 Nukleinsäuresequenz kodierend für Sojabohne (Glycine max) 2S Albumin (GenBank Acc.-No.: AF005030)
44. SEQ ID NO: 44  
Proteinsequenz kodierend für Sojabohne (Glycine max) 2S Albumin  
30
45. SEQ ID NO: 45  
Nukleinsäuresequenz kodierend für Brassica nigra 2S Speicherprotein (GenBank Acc.-No.: X65039)  
35
46. SEQ ID NO: 46  
Proteinsequenz kodierend für Brassica nigra 2S Speicherprotein
- 40 47. SEQ ID NO: 47  
Nukleinsäuresequenz kodierend für Sinapis alba sin5 Speicherprotein (GenBank Acc.-No.: X91798)
48. SEQ ID NO: 48  
45 Proteinsequenz kodierend für Sinapis alba sin5 Speicherprotein

## 54

49. SEQ ID NO: 49  
Nukleinsäuresequenz kodierend für Sonnenblume HaG5 2 S Albumin (GenBank Acc.-No.: X06410)
- 5 50. SEQ ID NO: 50  
proteinsequenz kodierend für Sonnenblume HaG5 2 S Albumin
51. SEQ ID NO: 51  
partielle Nukleinsäuresequenz kodierend für Sonnenblume (Helianthus annuus) 2S Albumin (GenBank Acc.-No.: X76101)
- 10
52. SEQ ID NO: 52  
partielle Proteinsequenz kodierend für Sonnenblume (Helianthus annuus) 2S Albumin
- 15
53. SEQ ID NO: 53  
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCru3 (Insert von Vektor pCR2.1-AtCRU3-RNAi)
- 20
54. SEQ ID NO: 54  
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCru3
- 25
55. SEQ ID NO: 55  
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCra1
- 30
56. SEQ ID NO: 56  
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCra1
- 35
57. SEQ ID NO: 57  
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 2S Speicherprotein At2S2
- 40
58. SEQ ID NO: 58  
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 2S Speicherprotein At2S2
- 45
59. SEQ ID NO: 59  
Nukleinsäuresequenz kodierend für Arabidopsis thaliana 12S Cruciferin Speicherprotein (ATCRU3; GenBank Acc.-No.: U66916)

## 55

60. SEQ ID NO: 60  
Proteinsequenz kodierend für Arabidopsis thaliana 12S Cruciferin Speicherprotein (ATCRU3)
- 5 61. SEQ ID NO: 61  
Nukleinsäuresequenz kodierend für A.thaliana 12S Speicherprotein (CRA1; GenBank Acc.-No.: M37247)
62. EQ ID NO: 62  
10 Proteinsequenz kodierend für A.thaliana 12S Speicherprotein (CRA1)
63. SEQ ID NO: 63  
Nukleinsäuresequenz kodierend für Arabidopsis thaliana 12S Speicherprotein AT5g44120/MLN1\_4 (GenBank Acc.-No.: AY070730)  
15
64. SEQ ID NO: 64  
Proteinsequenz kodierend für Arabidopsis thaliana 12S Speicherprotein AT5g44120/MLN1\_4  
20
65. SEQ ID NO: 65  
Nukleinsäuresequenz kodierend für Arabidopsis 12S Speicherprotein (CRB; GenBank Acc.-No.: X14313; M37248)
- 25 66. SEQ ID NO: 66  
Proteinsequenz kodierend für Arabidopsis 12S Speicherprotein (CRB)
67. SEQ ID NO: 67  
30 Nukleinsäuresequenz kodierend für Arabidopsis thaliana putatives 12S Speicherprotein (aus GenBank Acc.-No.: AC003027)
68. SEQ ID NO: 68  
Proteinsequenz kodierend für Arabidopsis thaliana putatives Speicherprotein (Protein\_id="AAD10679.1")  
35
69. SEQ ID NO: 69  
Nukleinsäuresequenz kodierend für Arabidopsis thaliana Cruciferin 12S Spwicherprotein (Atlg03890) (GenBank Acc.-No.: AY065432)  
40
70. SEQ ID NO: 70  
Proteinsequenz kodierend für Arabidopsis thaliana Cruciferin 12S Speicherprotein (Atlg03890)  
45



## 56

71. SEQ ID NO: 71  
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* Prohibitin 1 (Atphb1) (GenBank Acc.-No.: U66594)
- 5 72. SEQ ID NO: 72  
Proteinsequenz kodierend für *Arabidopsis thaliana* Prohibitin 1 (Atphb1)
- 10 73. SEQ ID NO: 73                      Oligonukleotidprimer OPN1
74. SEQ ID NO: 74                      Oligonukleotidprimer OPN2
75. SEQ ID NO: 75                      Oligonukleotidprimer OPN3
- 15 76. SEQ ID NO: 76                      Oligonukleotidprimer OPN4
77. SEQ ID NO: 77                      Oligonukleotidprimer OPN5
78. SEQ ID NO: 78                      Oligonukleotidprimer OPN6
- 20 79. SEQ ID NO: 79                      Oligonukleotidprimer OPN7
80. SEQ ID NO: 80                      Oligonukleotidprimer OPN8
- 25 81. SEQ ID NO: 81                      Oligonukleotidprimer OPN9
82. SEQ ID NO: 82                      Oligonukleotidprimer OPN10
83. SEQ ID NO: 83
- 30 Nukleinsäuresequenz kodierend für sRNAi4-dsRNA zur Suppression mehrerer Speicherproteine
84. SEQ ID NO: 84
- 35 Ribonukleinsäuresequenz kodierend für sRNAi4-dsRNA zur Suppression mehrerer Speicherproteine
85. SEQ ID NO: 85
- 40 Nukleinsäuresequenz kodierend für sRNAi8-dsRNA zur Suppression mehrerer Speicherproteine
86. SEQ ID NO: 86
- 45 Ribonukleinsäuresequenz kodierend für sRNAi8-dsRNA zur Suppression mehrerer Speicherproteine
87. SEQ ID NO: 87                      Oligonukleotidprimer OPN11

## 57

88. SEQ ID NO: 88                      Oligonukleotidprimer OP12
89. SEQ ID NO: 89                      Oligonukleotidprimer OPN13
- 5 90. SEQ ID NO: 90                      Oligonukleotidprimer OPN15
91. SEQ ID NO: 91                      Oligonukleotidprimer OPN16
92. SEQ ID NO: 92                      Oligonukleotidprimer OPN17
- 10 93. SEQ ID NO: 93  
Nukleinsäuresequenz kodierend für Arabidopsis thaliana "globulin-like protein" (GenBank Acc.-No.: NM\_119834)
- 15 94. SEQ ID NO: 94  
Proteinsequenz kodierend für Arabidopsis thaliana "globulin-like protein" (Protein\_id="NP\_195388.1)
95. SEQ ID NO: 95
- 20 Nukleinsäuresequenz kodierend für Glycine max 7S Samenglobulin (GenBank Acc.-No.: U59425)
96. SEQ ID NO: 96  
Proteinsequenz kodierend für für Glycine max 7S Samenglobulin
- 25 97. SEQ ID NO: 97  
Nukleinsäuresequenz kodierend für Zea mays 19kD Zein (GenBank Acc.-No.: E01144)
- 30 98. SEQ ID NO: 98  
Proteinsequenz kodierend für Zea mays 19kD Zein
99. SEQ ID NO: 99  
Nukleinsäuresequenz kodierend für Zea mays 19kD alpha Zein B1
- 35 (GenBank Acc.-No.: AF371269)
100. SEQ ID NO: 100  
Proteinsequenz kodierend für Zea mays 19kD alpha Zein B1
- 40 101. SEQ ID NO: 101  
Nukleinsäuresequenz kodierend für Zea mays 19kD alpha Zein B2 (GenBank Acc.-No.: AF371270)
102. SEQ ID NO: 102
- 45 Proteinsequenz kodierend für Zea mays 19kD alpha Zein B2

## 58

103. SEQ ID NO: 103  
Nukleinsäuresequenz kodierend für Zea mays 22kD alpha-zein  
(GenBank Acc.-No.: X61085)
- 5 104. SEQ ID NO: 104  
Proteinsequenz kodierend für Zea mays 22kD alpha-zein
105. SEQ ID NO: 105  
Nukleinsäuresequenz kodierend für Oryza sativa Prolamin  
10 (GenBank Acc.-No.: AB016503)
106. SEQ ID NO: 106  
Proteinsequenz kodierend für Oryza sativa Prolamin
- 15 107. SEQ ID NO: 107  
Nukleinsäuresequenz kodierend für A.sativa Avenin (GenBank  
Acc.-No.: M38446)
108. SEQ ID NO: 108  
20 Proteinsequenz kodierend für A.sativa Avenin
109. SEQ ID NO: 109  
Nukleinsäuresequenz kodierend für Hordeum vulgare C-Hordein  
(GenBank Acc.-No.: M36941)
- 25 110. SEQ ID NO: 110  
Proteinsequenz Teil 1 kodierend für Hordeum vulgare C-Hordein
111. SEQ ID NO: 111  
30 Proteinsequenz Teil 2 kodierend für Hordeum vulgare C-Hordein
112. SEQ ID NO: 112  
Nukleinsäuresequenz kodierend für Triticum aestivum LMW Glu-  
tenin-1D1 (GenBank Acc.-No.: X13306)
- 35 113. SEQ ID NO: 113  
Proteinsequenz kodierend für Triticum aestivum LMW Glute-  
nin-1D1
- 40 114. SEQ ID NO: 114  
Binärer Expressionsvektor für Agrobakterium vermittelte  
pflanzentransformation pSUN2-USP.

## 59

115. SEQ ID NO: 115  
Partielle Nukleinsäuresequenz kodierend für Homogentisat-1,2-dioxygenase aus *Brassica napus* (HGD; EC-Nr.: 1.13.11.5)
- 5
116. SEQ ID NO: 116  
Nukleinsäuresequenz kodierend für Homogentisat-1,2-dioxygenase aus *Arabidopsis thaliana* (HGD; EC-Nr.: 1.13.11.5)
- 10 117. SEQ ID NO: 117  
Proteinsequenz kodierend für Homogentisat-1,2-dioxygenase aus *Arabidopsis thaliana* (HGD; EC-Nr.: 1.13.11.5)
118. SEQ ID NO: 118  
15 Nukleinsäuresequenz kodierend für Maleylacetoacetatisomerase aus *Arabidopsis thaliana* (MAAI; EC-Nr.: 5.2.1.2.)
119. SEQ ID NO: 119  
Proteinsequenz kodierend für Maleylacetoacetatisomerase aus  
20 *Arabidopsis thaliana* (MAAI; EC-Nr.: 5.2.1.2.)
120. SEQ ID NO: 120  
Nukleinsäuresequenz kodierend für Fumarylacetoacetathydrolase aus *Arabidopsis thaliana* (FAAH; EC-Nr.: 3.7.1.2)
- 25
121. SEQ ID NO: 121  
Proteinsequenz kodierend für Fumarylacetoacetathydrolase aus *Arabidopsis thaliana* (FAAH; EC-Nr.: 3.7.1.2)
- 30 122. SEQ ID NO: 122  
Nukleinsäuresequenz kodierend für Suppressionskonstrukt 2 (p3300.1-Toc159-GFP-RNAi)
- 35 123. SEQ ID NO: 123                      Oligonukleotidprimer OPN18
124. SEQ ID NO: 124                      Oligonukleotidprimer OPN19
125. SEQ ID NO: 125                      Oligonukleotidprimer OPN20
- 40 126. SEQ ID NO: 126                      Oligonukleotidprimer OPN21

## Abbildungen

- 45 1. Fig.1: Schematische Darstellung der Speicherprotein-Suppressionskonstrukte.

## 60

Insert aus Vektor pCR2.1-sRNAi4 (1) (vgl. Beispiel 2d) und pCR2.1-sRNAi8 (2) (vgl. Beispiel 2e) kodierend für eine die AtCru3-, AtCRB und At2S3-Expression supprimierende dsRNA.

5 In den beiden Konstrukten sind die "sense"-Ribonukleotidsequenzen und die komplementären "antisense"-Ribonukleotidsequenzen (symbolisiert durch auf dem Kopf stehende Buchstaben) für die einzelnen zu supprimierenden Zielgene (AtCru3; AtCRB, At2S3) unterschiedlich angeordnet. Schraffierte Bereiche (I1, 10 I2 etc.) stellen Intronsequenzen (Linker) dar.

2. Fig.2A-D: Symbolische Darstellung verschiedener dsRNAs in ihrer Sekundärstruktur.

15 S1, S2, ... S(n): "sense"-Ribonukleotidsequenz  
AS1, AS2, ... AS(n): "antisense"-Ribonukleotidsequenz  
I: Intronsequenz

20 Die Anordnung der einzelnen "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen kann so erfolgen, dass zunächst alle "sense"-Ribonukleotidsequenzen aneinander gereiht werden und so quasi einen "sense"-Strang bilden, worauf dann alle "antisense"-Ribonukleotidsequenzen aneinander zu einem "antisense"-Strang zusammengefügt werden (A und C).

25 Alternativ kann die Anordnung der einzelnen "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen so erfolgen, dass Paare von jeweils komplementären "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen aneinander 30 gefügt werden (B und D).

"sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen können direkt aneinandergefügt sein (A und B) oder aber durch weitere Sequenzen beispielsweise Introns (I) von 35 einander getrennt sein (C und D).

3. Fig.3A-C: Symbolische Darstellung verschiedener dsRNAs in ihrer Sekundärstruktur.

40 S1, S2, ... S(n): "sense"-Ribonukleotidsequenz  
AS1, AS2, ... AS(n): "antisense"-Ribonukleotidsequenz  
SP: "SPACER"  
RE: Erkennungssequenz für Ribozym  
R: Ribozym

45

## 61

"sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen können durch weitere Sequenzen ("SPACER"; SP) voneinander getrennt sein (A). Der Spacer kann beispielsweise eines Erkennungssequenz für ein Ribozym sein. Das korrespondierende Ribozym kann separat exprimiert werden (B) oder aber auch ebenfalls von dem Spacer kodiert sein (C).

4. Fig.4: Abbildung des Suppressionskonstrukts mit den entsprechenden Restriktionsenzymchnittstellen:

5. Fig.5A: Identifikation einer Pflanze, die den Albino-Phänotyp zeigt (links). Der Phänotyp ist identisch zur ppi2 Mutante, die Toc159 nicht mehr exprimieren kann. Als Kontrolle Pflanzen mit Wildtyp Phänotyp, die parallel gewachsen sind.

Fig. 5B: Fluoreszenz-Analyse der Pflanzen aus Fig.5A. Anregung der Fluoreszenz durch Licht im Wellenlängenbereich 470-490 nm. Es wurde dieselbe Vergrößerung gewählt wie in Fig.5A.

## Beispiele

## Allgemeine Methoden:

- 5 Alle Chemikalien, wenn nicht anders erwähnt, stammen von den Firmen Fluka (Buchs), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deisenhofen). Restriktionsenzyme, DNA-modifizierende Enzyme und Molekularbiologie-Kits wurden von den Firmen Amersham-Pharmacia (Freiburg), Biometra (Göttingen), Roche  
10 (Mannheim), New England Biolabs (Schwalbach), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Qiagen (Hilden), Stratagen (Amsterdam, Niederlande), Invitrogen (Karlsruhe) und Ambion (Cambridgeshire, United Kingdom). Die verwendeten Reagenzien wurden entsprechend der Angaben des Herstellers eingesetzt.
- 15 Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die im Rahmen der vorliegenden Erfindung durchgeführten Klonierungsschritte wie z.B. Restriktionsspaltungen, Agarosegelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrozellulose und Nylonmembranen, Verknüpfen von DNA-Fragmenten, Transformation von *E. coli* Zellen, Anzucht von Bakterien, Vermehrung von Phagen und Sequenzanalyse rekombinanter DNA werden  
20 wie bei Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6 beschrieben durchgeführt. Die Sequenzierung rekombinanter DNA-Moleküle erfolgt mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma ABI nach der Methode von Sanger (Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-5467).
- 30

## Beispiel 1: Allgemeine Verfahren

- Die Pflanze *Arabidopsis thaliana* repräsentiert ein Mitglied der  
35 höheren Pflanzen (Samenpflanzen). Diese Pflanze ist eng verwandt mit anderen Pflanzenarten aus der Familie der Cruciferen wie z.B. *Brassica napus*, aber auch mit anderen Pflanzenfamilien der Dikotyledonen. Aufgrund des hohen Grades an Homologie ihrer DNA-Sequenzen bzw. Polypeptidsequenzen kann *Arabidopsis thaliana* als  
40 Modellpflanze für andere Pflanzenarten eingesetzt werden.

a) Anzucht von *Arabidopsis* Pflanzen

- Die Pflanzen werden entweder auf Murashige-Skoog Medium mit 0,5 %  
45 Saccharose (Ogas et al. (1997) Science 277:91-94) oder auf Erde gezogen (Focks & Benning (1998) Plant Physiol 118:91-101). Um einheitliche Keimungs- und Blühzeiten zu erreichen, werden die

## 63

Samen nach Ausplattieren bzw. Ausstreuen auf Erde zwei Tage bei 4° C stratifiziert. Nach der Blüte werden die Schoten markiert. Entsprechend der Markierungen werden dann Schoten mit einem Alter von 6 bis 20 Tagen nach der Blüte geerntet.

5

b) Isolierung von total RNA und poly-A<sup>+</sup> RNA aus Pflanzen

Für die Herstellung von Suppressionskonstrukten wird RNA bzw. polyA<sup>+</sup> RNA isoliert. RNA wurde aus Schoten von Arabidopsis Pflanzen nach folgender Vorschrift isoliert: Schotenmaterial im Alter von 6 bis 20 Tage nach Blüte wurde geerntet und in flüssigem Stickstoff schockgefroren. Das Material wurde vor der weiteren Verwendung bei -80°C gelagert. 75 mg des Materials wurde im gekühlten Mörser zu einem feinem Pulver gemahlen und mit 200 µL des Lysis-Puffers aus dem Ambion RNAqueös-Kit versetzt. Die Isolierung der totalen RNA wurde dann nach Herstellerangaben durchgeführt. Die RNA wurde mit 50 µL Elutionspuffer (Ambion) eluiert und die Konzentration durch Absorption einer 1 zu 100 verdünnten Lösung am Photometer (Eppendorf) bei 260 nm bestimmt. 40 µg/ml RNA entspricht dabei einer Absorption von 1. Die RNA-Lösungen wurden mit RNase freiem Wasser auf eine Konzentration von 1 µg/µL eingestellt. Die Konzentrationen wurden durch Agarosegelelektrophorese überprüft. Zur Isolierung von polyA<sup>+</sup> RNA wurde oligo(dT)-Zellulose von Amersham Pharmacia nach Herstellerangaben verwendet. RNA bzw. polyA<sup>+</sup> RNA wurde bei -70°C gelagert.

## c) Konstruktion der cDNA-Bank

Zur Konstruktion der cDNA-Bank aus Arabidopsis Schoten-RNA wurde die Erststrangsynthese unter Verwendung von Reverser Transkriptase aus Maus-Leukämie-Virus (Clontech) und Oligo-d(T)-Primern, die Zweitstrangsynthese durch Inkubation mit DNA-Polymerase I, Klenow-Enzym und RNase H-Spaltung bei 12°C (2 Std.), 16°C (1 Std.) und 22°C (1 Std.) erzielt. Die Reaktion wurde durch Inkubation bei 65°C (10 min) gestoppt und anschließend auf Eis überführt. Doppelsträngige DNA-Moleküle wurde mit T4-DNA-Polymerase (Roche, Mannheim) bei 37°C (30 min) mit glatten Enden versehen. Die Nukleotide wurden durch Phenol/Chloroform-Extraktion und Sephadex-G50-Zentrifugiersäulen entfernt. EcoRI/XhoI-Adapter (Pharmacia, Freiburg, Deutschland) wurden mittels T4-DNA-Ligase (Roche, 12°C, über Nacht) an die cDNA-Enden ligiert, mit XhoI nachgeschnitten und durch Inkubation mit Polynukleotidkinase (Roche, 37°C, 30 min) phosphoryliert. Dieses Gemisch wurde der Trennung auf einem Low-Melting-Agarose-Gel unterworfen. DNA-Moleküle über 300 Basenpaaren wurden aus dem Gel eluiert, Phenol-extrahiert, auf Elutip-D-Säulen (Schleicher und Schüll, Dassel, Deutschland) konzentriert und an Vektorarme ligiert und in lambda-ZAPII-Phagen



oder lambda-ZAP-Express-Phagen unter Verwendung des Gigapack Gold-Kits (Stratagene, Amsterdam, Niederlande) verpackt, wobei Material des Herstellers verwendet und seine Anweisungen befolgt wurden.

5

- d) Isolierung von genomischer DNA aus Pflanzen wie *Arabidopsis thaliana* oder *Brassica napus* (CTAB-Methode)

Zur Isolierung genomischer DNA aus Pflanzen wie *Arabidopsis thaliana* oder *Brassica napus* werden ca. 0,25 g Blattmaterial junger Pflanzen im vegetativen Stadium in flüssigem Stickstoff zu feinem Pulver gemörserst. Das pulverisierte Pflanzenmaterial wird zusammen mit 1 ml 65°C-warmem CTAB I-Puffer (CTAB: Hexadecyltrimethylammoniumbromid, auch genannt Cetyltrimethylammoniumbromid; Sigma Kat.-Nr.: H6269) und 20 µl β-Mercaptoethanol in einen vorgewärmten zweiten Mörser gegeben und nach vollständiger Homogenisierung wird der Extrakt in ein 2 ml Eppendorf-Gefäß überführt und für 1 h bei 65°C unter regelmäßiger, vorsichtiger Durchmischung inkubiert. Nach Abkühlung auf Raumtemperatur wird der Ansatz mit 1 ml Chloroform/Octanol (24:1, mit 1M Tris/HCl, pH 8,0 ausgeschüttelt) durch langsames Invertieren extrahiert und zur Phasentrennung für 5 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Anschließend wird die wässrige Phase erneut mit 1 ml Chloroform/Octanol extrahiert, zentrifugiert und durch Invertieren mit 1/10 Volumen auf 65°C vorgewärmtem CTAB II-Puffer sorgfältig gemischt. Anschließend wird der Ansatz durch vorsichtiges Schwenken mit 1 ml Chloroform/Octanol-Gemisch (siehe oben) versetzt und zur erneuten Phasentrennung für 5 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Die wässrige untere Phase wird in ein frisches Eppendorf-Gefäß überführt und die obere organische Phase wird in einem frischen Eppendorf-Gefäß erneut für 15 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Die hieraus resultierende wässrige Phase wird mit der wässrigen Phase des vorherigen Zentrifugationsschrittes vereinigt und der gesamte Ansatz mit exakt demselben Volumen vorgewärmtem CTAB III-Puffer versetzt. Es folgt eine Inkubation bei 65°C, bis die DNA in Flocken ausfällt. Dies kann bis zu 1 h dauern oder durch Inkubation bei 37°C über Nacht erfolgen. Das aus dem anschließenden Zentrifugationsschritt (5 min, 2000 rpm (500 x g), 4°C) resultierende Sediment wird mit 250 µl auf 65°C vorgewärmtem CTAB IV-Puffer versetzt und für mindestens 30 min bzw. bis zur vollständigen Auflösung des Sediments bei 65°C inkubiert. Anschließend wird die Lösung zur Fällung der DNA mit 2,5 Volumina eiskaltem Ethanol vermischt und für 1h bei -20°C inkubiert. Alternativ kann der Ansatz mit 0.6 Volumina Isopropanol vermischt und ohne weitere Inkubation sofort für 15 min bei 8,500 rpm (7,500 x g) und 4°C zentrifugiert werden. Die sedimentierte DNA wird durch

## 65

Invertieren des Eppendorf-Gefäßes zweimal mit je 1 ml 80%igem eiskaltem Ethanol gewaschen, nach jedem Waschschrift erneut zentrifugiert (5 min, 8,500 rpm (7,500 x g), 4°C) und anschließend für ca. 15 min luftgetrocknet. Abschließend wird die DNA in  
5 100 µl TE mit 100 µg/ml RNase resuspendiert und für 30 min bei Raumtemperatur inkubiert. Die DNA Lösung ist nach einer weiteren Inkubationsphase über Nacht bei 4°C homogen und kann für weiterführende Experimente verwendet werden.

## 10 Lösungen für CTAB:

Lösung I (für 200 ml):

- 100 mM Tris/HCl pH 8,0 (2,42 g)
- 1,4 M NaCl (16,36 g)
- 15 20 mM EDTA (8,0 ml von 0,5 M Stammlösung)
- 2 % (w/v) CTAB (4,0 g)

Jeweils vor der Verwendung werden frisch zugesetzt:  
2 % β-Mercaptoethanol (20 µl für 1 ml Lösung I).

20

Lösung II (für 200 ml):

- 0,7 M NaCl (8,18 g)
- 10 % (w/v) CTAB (20 g)

## 25 Lösung III (für 200 ml):

- 50 mM Tris/HCl pH 8,0 (1,21 g)
- 10 mM EDTA (4 ml 0,5 M von 0,5 M Stammlösung)
- 1 % (w/v) CTAB (2,0 g)

## 30 Lösung IV (High-salt TE) (für 200 ml):

- 10 mM Tris/ HCl pH 8,0 (0,242 g)
- 0,1 mM EDTA (40 µl 0,5 M Stammlösung)
- 1 M NaCl (11, 69 g)

## 35 Chloroform/Octanol (24:1) (für 200 ml):

- 192 ml Chloroform
- 8 ml Octanol

Die Mischung wird 2x mit 1 M TrisHCl pH 8,0 ausgeschüttelt und vor Licht geschützt gelagert.

40

## Beispiel 2: Herstellung von Suppressionskonstrukten

- Ausgehend von der genomischer Arabidopsis thaliana DNA oder cDNA wurden über PCR mittels der aufgeführten Oligonukleotide folgende  
45 Fragmente von Speicherproteinsequenzen amplifiziert. Dabei kam nachfolgendes PCR Protokoll zum Einsatz:

## 66

Zusammensetzung des PCR-Ansatzes (50 µL):

- 5,00 µL Template cDNA oder genomische DNA (ca. 1 µg)
- 5,00 µL 10x Puffer (Advantage-Polymerase) + 25 mM MgCl<sub>2</sub>
- 5    5,00 µL 2mM dNTP
- 1,25 µL je Primer (10 pmol/µL)
- 0,50 µL Advantage-Polymerase (Clontech)

PCR-Programm: Anfangsdenaturierung für 2 min bei 95°C, dann 35 Zyklen mit 45 sec 95°C, 45 sec 55°C und 2 min 72°C. Abschliessende Extension von 5 min bei 72°C.

a) Ausgangsvektor pCR2.1-AtCRU3-RNAi

- 15 Aus genomischer Arabidopsis thaliana DNA wird mit nachfolgendem Oligonukleotid-Primerpaar ein Exonbereich mit dem vollständigen anschließenden Intron einschliesslich der an das Intron anschließenden Spleiß-Akzeptorsequenz des 12S Speicherprotein AtCRU3 (Basenpaar 1947 bis 2603 der Sequenz mit der GenBank Acc.-No: U66916) amplifiziert:

ONP1 (SEQ ID NO: 134):

5'-ATAAGAATGCGGCCGCGTGTTCATTTGGCCGGAACAAC-3'

- 25 ONP2 (SEQ ID NO: 135):

5'-CCCGGATCCTTCTGTAAACATTTGACAAAACATG-3'

- Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-1 Vektor und die Sequenz überprüft.

- Für die den antisense-Strang der dsRNA kodierende Sequenz wird aus Arabidopsis thaliana cDNA lediglich das gleiche Exon wie oben (Basenpaar 1947 bis 2384) mit dem nachfolgenden Primerpaar amplifiziert:

ONP3 (SEQ ID NO: 136):

5'-ATAAGAATGCGGCCGCGTGTTCATTTGGCCGGAACAAC-3'

- 40 ONP4 (SEQ ID NO: 137):

5' ATAAGAATGCGGCCGCGGATCCACCCTGGAGAACGCCACGAGTG-3'

- Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-2 Vektor und die Sequenz überprüft.

## 67

0,5 µg von Vektor pCR2.1-1 werden mit dem Restriktionsenzym BamHI (New England Biolabs) für 2 Stunden nach Herstellerangaben inkubiert und dann für 15 min mit alkalischer Phosphatase (New England Biolabs) dephosphoryliert. Der so präparierte Vektor (1 µL) wird dann mit dem aus Vektor pCR2.1-2 gewonnenen Fragment ligiert. Dazu werden 0,5 µg von Vektor pCR2.1-2 2 Stunden mit BamHI (New England Biolabs) verdaut und die DNA-Fragmente per Gelelektrophorese aufgetrennt. Das neben dem Vektor (3,9 kb) entstandene 489 bp grosse Stück wird aus dem Gel ausgeschnitten und mit dem "Gelpurification"-Kit (Qiagen) nach Herstellerangaben aufgereinigt und mit 50 µL Elutionspuffer eluiert. 10 µL des Eluates werden mit Vektor pCR2.1-1 (s.o.) über Nacht bei 16°C ligiert (T4 Ligase, New England Biolabs). Die Ligationsprodukte werden dann in TOP10 Zellen (Stratagene) nach Herstellerangaben transformiert und entsprechend selektioniert. Positive Klone werden mit dem Primerpaar ONP1 und ONP2 durch PCR verifiziert. Der erhaltene Vektor wird pCR2.1-AtCRU3-RNAi genannt. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 105 beschrieben.

## 20 b) Ausgangsvektor pCR2.1-AtCRB-RNAi

Mit nachfolgendem Oligonukleotid-Primerpaar wird ein Exonbereich des 12S Speicherprotein AtCRB (SEQ ID NO: 117 bzw. 118; Basenpaar 601 bis 1874 der Sequenz mit der GenBank Acc.-No: M37248) aus Arabidopsis thaliana cDNA amplifiziert:

ONP5 (SEQ ID NO: 138):

5' ATAAGAATGCGGCCGCGGATCCCTCAGGGTCTTTTCTTGCCCACT-3'

30 ONP6 (SEQ ID NO: 139):

5'-CCGCTCGAGTTTACGGATGGAGCCACGAAG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-3 Vektor und die Sequenz überprüft.

Für den als Linker fungierenden Bereich wird aus Arabidopsis thaliana genomischer DNA ein Intron mit den entsprechenden Spliceakzeptor und -donorsequenzen der flankierenden Exons (Basenpaar 1874 bis 2117 der Sequenz mit der GenBank Acc.-No: M37248) mit dem nachfolgenden Primerpaar amplifiziert:

ONP7 (SEQ ID NO: 140):

5'-CCGCTCGAGGTAAGCTCAACAAATCTTTAG-3'

45 ONP8 (SEQ ID NO: 141):

5'-ACGCGTCGACGCGTTCTGCGTGCAAGATATT-3'

## 68

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-4 Vektor und die Sequenz überprüft.

- 5 Das Konstrukt für AtCRB wird in einer ähnlichen Strategie wie für AtCRU3 erläutert, erstellt. Vektor pCR2.1-3 wird mit mit XhoI (New England Biolabs) für 2 Stunden inkubiert und dephosphoryliert (alkalische Phosphatase, New England Biolabs). Vektor pCR2.1-4 wird ebenfalls mit XhoI in derselben Weise inkubiert und
- 10 die Gelfragmente per Gelelektrophorese aufgetrennt. Die entsprechenden Fragmente werden in der unter AtCRU3 beschriebenen Art und Weise aufgereinigt und ligiert, resultierend nach Bakterientransformation in dem Vektor pCR2.1-AtCRB Exon/Intron. Dieser Vektor wird für 2 Stunden mit XbaI (NEB), anschliessend für 15
- 15 min mit Klenow-Fragment (NEB), dann für 2 Stunden mit SalI inkubiert und zuletzt 15 min mit alkalischer Phosphatase (NEB) behandelt. Parallel wird der Vektor pCR2.1-3 mit BamHI (NEB), dann 15 min mit Klenow-Fragment und anschliessend 2 Stunden mit XhoI (NEB) inkubiert. Das Exon-Fragment von AtCRB wird nach Gelelektrophorese isoliert, gereinigt und zur Ligation eingesetzt. Beide
- 20 Fragmente wurden dann ligiert und der Vektor pCR2.1-AtCRB-RNAi resultierte. Der erhaltene Vektor wird pCR2.1-AtCRB-RNAi genannt. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 107 beschrieben.

- 25 c) Ausgangsvektor pCR2.1-At2S3-RNAi

Mit nachfolgendem Oligonukleotid-Primerpaar wird ein Exonbereich des 2S Speicherprotein At2S3 (SEQ ID NO: 3 bzw. 4; Basenpaar 212 bis 706 der Sequenz mit der GenBank Acc.-No: M22035) amplifiziert:

30

ONP9 (SEQ ID NO: 142):

5'-ATAAGAATGCGGCCGCGATCCATGGCTAACAAGCTCTTCCTCGTC-3'

- 35 ONP10 (SEQ ID NO: 143):

5'-ATAAGAATGCGGCCGCGATCCCTAGTAGTAAGGAGGGAAGAAAG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-5

40 Vektor und die Sequenz überprüft. Für den als Linker fungierenden Bereich wird das gleiche Intron wie unter b) mit den Primern OPN 7 und OPN 8 amplifiziert eingesetzt.

- Das Konstrukt für At2S3 wird in einer ähnlichen Strategie wie für
- 45 AtCRU3 erläutert, erstellt. Vektor pCR2.1-5 wird mit mit XhoI (New England Biolabs) für 2 Stunden inkubiert und dephosphoryliert (alkalische Phosphatase, New England Biolabs). Vektor

pCR2.1-3 werden ebenfalls mit XhoI in derselben Weise inkubiert und die Gelfragmente per Gelelektrophorese aufgetrennt. Die entsprechenden Fragmente werden in der unter AtCRU3 beschriebenen Art und Weise aufgereinigt und ligiert, resultierend nach Bakterientransformation in dem Vektor pCR2.1-At2S3 Exon/Intron. Dieser Vektor wird für 2 Stunden mit SalI (NEB), anschliessend für 15 min mit Klenow-Fragment (NEB) inkubiert und zuletzt 15 min mit alkalischer Phosphatase (NEB) behandelt. Parallel wird der Vektor pCR2.1-5 mit BamHI (NEB) und dann 15 min mit Klenow-Fragment inkubiert. Das Exon-Fragment von At2S3 wird nach Gelelektrophorese isoliert, gereinigt und zur Ligation eingesetzt. Beide Fragmente werden dann ligiert und der Vektor pCR2.1-At2S3-RNAi resultierte. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 109 beschrieben.

15

d) Herstellung von Super-Suppressionskonstrukt 1

- Die Vektoren pCR2.1-AtCRU3-RNAi und pCR2.1-4 (siehe oben) werden mit den Restriktionsenzymen XhoI und SalI für 2 Stunden bei 37°C inkubiert, die DNA-Fragmente durch Agarose-Gelelektrophorese aufgetrennt und sowohl der Vektor als auch das PCR-Insert aus pCR2.1-4 ausgeschnitten und mit dem "Gelpurification"-Kit von Qiagen nach Herstellerangaben aufgereinigt und mit 50 µL Elutionspuffer eluiert. Vom Vektor wird 1 µL, vom PCR-Insert aus pCR2.1-4 8 µL der Eluate für die Ligation eingesetzt, resultierend in dem Konstrukt pCR2.1-sRNAi1. Dieser Vektor wird für 2 Stunden mit dem Restriktionsenzym XhoI und dann für 15 min mit Klenow-Fragment inkubiert.
- Der Vektor pCR2.1-AtCRB-RNAi (siehe oben) wird mit dem Enzym EcoRI für 2 Stunden inkubiert und ebenfalls 15 min mit Klenow-Fragment behandelt. Beide Inkubationsansätze werden durch Gelelektrophorese aufgetrennt und jeweils der Vektor (pCR2.1-sRNAi1) bzw. das Insert (aus pCR2.1-AtCRB-RNAi) aus dem Agarosegel ausgeschnitten und die DNA-Fragmente wie oben beschrieben aufgereinigt. Für die Ligation werden 1 µL des Eluates vom Vektor und 8 µL des Eluates vom Insert eingesetzt und bei 4°C über Nacht inkubiert. Das resultierende Konstrukt wird mit pCR2.1-sRNAi2 bezeichnet. Der resultierende Vektor wird mit dem Enzym XbaI und anschliessend mit Klenow-Fragment inkubiert. Der Vektor pCR2.1-4 wird mit den Enzymen EcoRV und XbaI und anschliessend mit Klenow-Fragment inkubiert. Nach Gelelektrophorese und -reinigung wird das Fragment aus pCR2.1-4 mit dem Vektor pCR2.1-sRNAi2 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi3. Der resultierende Vektor wird dann mit dem Enzym ApaI für 2 Stunden und dann mit Klenow-Fragment für 15 min inkubiert. Als Insert wird der Vektor pCR2.1-At2S3-RNAi mit dem Enzym EcoRI für 2 Stunden und dann mit

Klenow-Fragment für 15 min inkubiert. Nach Gelelektrophorese und -reinigung werden die Eluate ligiert, resultierend in dem Vektor pCR2.1-sRNAi4. Aus diesem Vektor wird dann das sRNAi4-Fragment (SEQ ID NO: 144; vgl. Fig. 1(1)), kodierend für die super-supprimierende dsRNA, durch Inkubation mit HindIII und PvuI ausgeschnitten und in den binären Vektor pSUN-USP (SEQ ID NO: 179) ligiert. Das Konstrukt dient der gleichzeitigen Suppression von Arabidopsis thaliana Speicherproteinen CRB (SEQ ID NO: 4), CRU3 (SEQ ID NO: 112) und At2S3 (SEQ ID NO: 118).

10

Der verwendete Vektor pSUN-USP ist ein binärer Vektor zur Pflanzentransformation auf Basis von pBinAR (Höfgen und Willmitzer (1990) Plant Science 66: 221-230). Eine gewebespezifische Expression im Samen lässt sich unter Verwendung des gewebespezifischen USP-Promotors USP-Promotors erzielt.

15

e) Herstellung von Super-Suppressionskonstrukt 2

Ausgehend von Arabidopsis thaliana cDNA wird ein Fragment aus dem Speicherprotein AtCRU3 (SEQ ID NO: 111, 112) mit dem nachfolgenden Oligonukleotid-Primerpaar unter den in Beispiel 2 angegebenen PCR-Bedingungen amplifiziert:

20

OPN 11: 5'-AAAAGGCCTGTGTTCCATTTGGCCGGAACAAC-3' (SEQ ID NO: 148)  
 25 OPN 12: 5'-AAAGATATCACCTGGAGAACGCCACGAGTG-3' (SEQ ID NO: 149).

Das erhaltene Fragment wird in den Vektor pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in den pCR2.1-6 und die Sequenzen überprüft.

30

Ausgehend von Arabidopsis thaliana cDNA wird ein Fragment aus dem Speicherprotein At2S3 (SEQ ID NO: 3, 4) mit dem nachfolgenden Oligonukleotid-Primerpaar unter den in Beispiel 2 angegebenen PCR-Bedingungen amplifiziert:

35

OPN 13: 5'-AAAAGGCCTATGGCTAACAAGCTCTTCCTCGTC-3' (SEQ ID NO: 150)  
 OPN 14: 5'-AAAGATATCCTAGTAGTAAGGAGGGAAGAAAG-3' (SEQ ID NO: 151).

Das erhaltene Fragment wird in den Vektor pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in den pCR2.1-7 und die Sequenzen überprüft.

40

Aus den pCR2.1-3, pCR2.1-4 (siehe Beispiel 2) und pCR2.1-6 und pCR2.1-7 werden dann die Konstrukte folgendermassen zusammen ligiert: Der Vektor pCR2.1-3 wird 2 Stunden mit EcoRV inkubiert und anschliessend 15 min mit alkalischer Phosphatase dephosphoryliert. Der Vektor pCR2.1-6 wird mit den Enzymen StuI und EcoRV

45

## 71

für 2 Stunden inkubiert und das PCR-Insert über Gelelektrophorese und -reinigung isoliert. Vektor pCR2.1-3 und Insert aus pCR2.1-6 werden dann über Nacht bei 4°C ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi5. Dieser Vektor wird dann mit EcoRV inkubiert und dephosphoryliert und mit dem StuI/ EcoRV inkubierten und gelaufgereinigten Fragment aus pCR2.1-7 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi6. Dieser Vektor wird dann mit XhoI inkubiert und dephosphoryliert. Der Vektor pCR2.1-4 wird mit SalI und XhoI inkubiert und das Insert aus pCR2.1-4 mit dem vorbereiteten Vektor pCR2.1-sRNAi6 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi7. Ausgehend von pCR2.1-sRNAi7 wird eine PCR mit den nachfolgenden Primerpaar unter den in Beispiel 2 gegebenen Bedingungen durchgeführt:

15 OPN 15: 5' CCGCTCGAGCTCAGGGTCTTTCTTGCCCACT (SEQ ID NO: 152)  
 OPN 16: 5'-CCGGTCGACCTAGTAGTAAGGAGGAAGAAAG (SEQ ID NO: 153).

Das resultierende PCR-Produkt wird mit den Enzymen XhoI und SalI inkubiert. Das Fragment wird dann in den Vektor pCR2.1-sRNAi7 (inkubiert mit XhoI) ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi8. Aus diesem Vektor wird dann das sRNAi8-Fragment (SEQ ID NO: 146; vgl. Fig. 1(2)), kodierend für die super-supprimierende dsRNA, durch Inkubation mit HindIII und XbaI geschnitten und in den binären Vektor pSUN-USP (SEQ ID NO: 179) ligiert. Das Konstrukt dient der gleichzeitigen Suppression von Arabidopsis thaliana Speicherproteinen CRB (SEQ ID NO: 4), CRU3 (SEQ ID NO: 112) und At2S3 (SEQ ID NO: 118).

## Beispiel 3: Transformation von Agrobacterium

30 Die Agrobacterium-vermittelte Pflanzentransformation kann zum Beispiel unter Verwendung der Agrobacterium tumefaciens-Stämme GV3101 (pMP90) (Koncz und Schell (1986) Mol Gen Genet 204: 383-396) oder LBA4404 (Clontech) durchgeführt werden. Die Transformation kann durch Standard-Transformationstechniken durchgeführt werden (Deblaere et al. (1984) Nucl Acids Res 13:4777-4788).

## Beispiel 4: Pflanzentransformation

40 Die Agrobacterium-vermittelte Pflanzentransformation kann unter Verwendung von Standard-Transformations- und Regenerationstechniken durchgeführt werden (Gelvin, Stanton B., Schilperoort, Robert A., Plant Molecular Biology Manual, 2. Aufl., Dordrecht: Kluwer Academic Publ., 1995, in Sect., Ringbuc Zentrale Signatur: BT11-P  
 45 ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., Methods



in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

Die Transformation mittels *Agrobacterium* von *Arabidopsis thaliana* wird durch die Methode nach Bechthold et al., 1993 (C.R. Acad. Sci. Ser. III Sci. Vie., 316, 1194-1199) durchgeführt. Beispielsweise kann Raps mittels Kotyledonen- oder Hypokotyltransformation transformiert werden (Moloney et al., Plant Cell Report 8 (1989) 238-242; De Block et al., Plant Physiol. 91 (1989) 694-701). Die Verwendung von Antibiotika für die *Agrobacterium*- und Pflanzenselektion hängt von dem für die Transformation verwendeten binären Vektor und *Agrobacterium*-Stamm ab. Die Rapsselektion wird gewöhnlich unter Verwendung von Kanamycin als selektierbarem Pflanzenmarker durchgeführt.

Der *Agrobacterium*-vermittelte Gentransfer in Lein (*Linum usitatissimum*) läßt sich unter Verwendung von beispielsweise einer von Mlynarova et al. (1994) Plant Cell Report 13:282-285 beschriebenen Technik durchführen.

Die Transformation von Soja kann unter Verwendung von beispielsweise einer in EP-A-0 0424 047 (Pioneer Hi-Bred International) oder in EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo) beschriebenen Technik durchgeführt werden.

Die Pflanzentransformation unter Verwendung von Teilchenschuß, Polyethylenglycol-vermittelter DNA-Aufnahme oder über die Siliziumcarbonatfaser-Technik ist beispielsweise beschrieben von Freeling und Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

Beispiel 5: Untersuchung der Expression eines rekombinanten Genproduktes in einem transformierten Organismus

Die Aktivität eines rekombinanten Genproduktes im transformierten Wirtsorganismus wurde auf der Transkriptions- und/oder der Translationsebene gemessen.

Ein geeignetes Verfahren zur Bestimmung der Menge an Transkription des Gens (ein Hinweis auf die Menge an RNA, die für die Translation des Genproduktes zur Verfügung steht) ist die Durchführung eines Northern-Blots wie unten ausgeführt (als Bezugsstelle siehe Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York, oder den oben erwähnten Beispielteil), wobei ein Primer, der so gestaltet ist, daß er an das Gen von Interesse bindet, mit einer nachweisbaren Markierung (gewöhnlich radioaktiv oder chemilumineszent) markiert wird, so daß, wenn die

## 73

Gesamt-RNA einer Kultur des Organismus extrahiert, auf einem Gel aufgetrennt, auf eine stabile Matrix transferiert und mit dieser Sonde inkubiert wird, die Bindung und das Ausmaß der Bindung der Sonde das Vorliegen und auch die Menge der mRNA für dieses Gen anzeigt. Diese Information zeigt den Grad der Transkription des transformierten Gens an. Zelluläre Gesamt-RNA kann aus Zellen, Geweben oder Organen mit mehreren Verfahren, die alle im Fachgebiet bekannt sind, wie zum Beispiel das von Bormann, E.R., et al. (1992) Mol. Microbiol. 6:317-326 beschriebene, präpariert werden.

10

## Northern-Hybridisierung:

Für die RNA-Hybridisierung wurden 20 µg Gesamt-RNA oder 1 µg poly(A)<sup>+</sup>-RNA mittels Gelelektrophorese in Agarosegelen mit einer Stärke von 1,25 % unter Verwendung von Formaldehyd, wie beschrieben in Amasino (1986, Anal. Biochem. 152, 304) aufgetrennt, mittels Kapillaranziehung unter Verwendung von 10 x SSC auf positiv geladene Nylonmembranen (Hybond N+, Amersham, Braunschweig) übertragen, mittels UV-Licht immobilisiert und 3 Stunden bei 68°C unter Verwendung von Hybridisierungspuffer (10 % Dextransulfat Gew./Vol., 1 M NaCl, 1 % SDS, 100 mg Heringssperma-DNA) vorhybridisiert. Die Markierung der DNA-Sonde mit dem Highprime DNA labeling-Kit (Roche, Mannheim, Deutschland) erfolgte während der Vorhybridisierung unter Verwendung von alpha-<sup>32</sup>P-dCTP (Amersham Pharmacia, Braunschweig, Deutschland). Die Hybridisierung wurde nach Zugabe der markierten DNA-Sonde im gleichen Puffer bei 68°C über Nacht durchgeführt. Die Waschschrte wurden zweimal für 15 min unter Verwendung von 2 X SSC und zweimal für 30 min unter Verwendung von 1 X SSC, 1 % SDS, bei 68°C durchgeführt. Die Exposition der verschlossenen Filter wurde bei -70°C für einen Zeitraum von 1 bis 14 T durchgeführt.

Zur Untersuchung des Vorliegens oder der relativen Menge an von dieser mRNA translatiertem Protein können Standardtechniken, wie ein Western-Blot, eingesetzt werden (siehe beispielsweise Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York). Bei diesem Verfahren werden die zellulären Gesamt-Proteine extrahiert, mittels Gelelektrophorese aufgetrennt, auf eine Matrix, wie Nitrozellulose, übertragen und mit einer Sonde, wie einem Antikörper, der spezifisch an das gewünschte Protein bindet, inkubiert. Diese Sonde ist gewöhnlich mit einer chemilumineszenten oder kolorimetrischen Markierung versehen, die sich leicht nachweisen läßt. Das Vorliegen und die Menge der beobachteten Markierung zeigt das Vorliegen und die Menge des gewünschten, in der Zelle vorliegenden mutierten Proteins an.

45

Beispiel 6: Analyse der Auswirkung der rekombinanten Proteine auf die Produktion des gewünschten Produktes

Die Auswirkung der genetischen Modifikation in Pflanzen, Pilzen,  
5 Algen, Ciliaten oder auf die Produktion einer gewünschten Verbindung (wie einer Fettsäure) kann bestimmt werden, indem die modifizierte Mikroorganismen oder die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen) gezüchtet werden und das Medium und/oder die zellulären Komponenten auf die  
10 erhöhte Produktion des gewünschten Produktes (d.h. von Lipiden oder einer Fettsäure) untersucht wird. Diese Analysetechniken sind dem Fachmann bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie (siehe  
15 beispielsweise Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et  
20 al. (1993) Biotechnology, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley  
25 and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988) Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

30

Neben den oben erwähnten Verfahren werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145, beschrieben extrahiert. Die qualitative  
35 und quantitative Lipid- oder Fettsäureanalyse ist beschrieben bei Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily  
40 Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

Zusätzlich zur Messung des Endproduktes der Fermentation ist  
45 es auch möglich, andere Komponenten der Stoffwechselwege zu analysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamteffizienz

## 75

- der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (z.B. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere Ionen), Messungen der Biomassezusammensetzung und des Wachstums,
- 5 Analyse der Produktion üblicher Metabolite von Biosynthesewegen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F. Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192
- 10 (ISBN: 0199635773) und darin angegebenen Literaturstellen beschrieben.

- Ein Beispiel ist die Analyse von Fettsäuren (Abkürzungen: FAME, Fettsäuremethylester; GC-MS, Gas-Flüssigkeitschromatographie-Massenspektrometrie; TAG, Triacylglycerin; TLC, Dünnschichtchromatographie).
- 15

- Der unzweideutige Nachweis für das Vorliegen von Fettsäureprodukten kann mittels Analyse rekombinanter Organismen nach Standard-
- 20 Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: Advances on Lipid Methodology, Vierte Aufl.: Christie, Oily Press, Dundee, 119-169; 1998, Gaschromatographie-Massenspektrometrie-Verfahren, Lipide 33:343-353).

25

- Das zu analysierende Material kann durch Ultraschallbehandlung, Mahlen in der Glasmühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren aufgebrochen werden. Das Material muss nach dem Aufbrechen zentrifugiert werden. Das Sediment
- 30 wird in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit 2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl- und Lipidverbindungen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester werden in Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapillarsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 Mikrom, 0,32 mm) bei einem Temperaturgradienten zwischen 170°C und 240°C für 20 min und 5 min bei 240°C unterworfen. Die Identität der erhaltenen Fettsäuremethylester muss unter Verwendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma), definiert werden.
- 40

Für die Öl-Analyse der mit den Suppressionskonstrukten transformierten Arabidopsis Pflanzen wird folgendes Protokoll angewendet:

45

## 76

Die Extraktion der Lipide aus Samen wird nach der Methode von Bligh & Dyer (1959) Can J Biochem Physiol 37:911 durchgeführt. Dazu werden 5 mg Arabidopsis Samen in 1,2 ml Qiagen-Microtubes (Qiagen, Hilden) auf einer Sartorius (Göttingen) Mikrowaage abgewogen. Das Samenmaterial wird mit 500 uL Chloroform/Methanol (2:1; enthält Mono-C17-glycerin von Sigma als internen Standard) in der Rätשמühle MM300 der Firma Retsch (Haan) homogenisiert und 20 min bei RT inkubiert. Nach Zugabe von 500 uL 50 mM Kaliumphosphatpuffer pH 7,5 erfolgt die Phasentrennung. Von der organischen Phase werden 50 µL abgenommen, mit 1500 uL Chloroform verdünnt und 5 µL auf die Kapillaren Chromarods SIII der Firma Iatroscan (SKS, Bechenheim) aufgetragen. Nach Auftrag der Proben werden diese für 15 min in einer Dünnschichtkammer, die gesättigt ist mit 6:2:2 Chloroform: Methanol: Toluol in einem ersten Schritt aufgetrennt. Nach Ablauf der Zeit werden die Kapillaren 4 min bei Raumtemperatur getrocknet und dann für 22 min in eine Dünnschichtkammer, die gesättigt ist mit 7:3 n-Hexan:Diethylether gestellt. Nach einem weiteren Trocknungsschritt für 4 min bei Raumtemperatur werden die Proben in einem Iatroscan MK-5 (SKS, Bechenheim) entsprechend Fraser & Taggart, 1988 J. Chromatogr. 439:404 analysiert. Folgende Parameter wurden für die Messungen eingestellt: Slice width 50 msec, Treshold 20 mV, Noise 30, Skim ratio 0. Die Quantifizierung der Daten erfolgte anhand des internen Standards Mono-C17-glycerin (Sigma) sowie einer erstellten Eichkurve mit Tri-C17-glycerin (Sigma) mittels des Programms ChromStar (SKS, Bechenheim).

Für die quantitative Bestimmung der Ölgehalte werden Samen von jeweils 10 Pflanzen derselben unabhängigen transgenen Linie analysiert. Insgesamt wurde der Ölgehalt von 30 transgene Linien der T1 Generation, 10 transgene Linien mit je 10 Pflanzen der T2 Generation und 5 transgene Linien mit je 10 Pflanzen der T3 Linien bestimmt. Dabei zeigen die transgenen Pflanzen einen signifikant höheren Ölgehalt als entsprechend gleichbehandelte Kontrollpflanzen.

## Beispiel 7:

Zum Nachweis der Funktionalität der multiplen RNAi Konstrukte wurden Gene ausgewählt, deren Supression einen deutlichen phänologischen Effekt hervorrufen. Ein solches Gen ist zum Beispiel Toc159. Dieses Gen ist essentiell für die Entwicklung und Funktionalität von Chloroplasten in Arabidopsis (Bauer et al. Nature, 403, 203-207). Ein Ausschalten dieses Gens führt zu chlorophylldefizienten Pflanzen, deren Blatt-Erscheinungsbild dann hell-grün bis weiss ist. Dieser Albino-Phänotyp ist sehr leicht zu unterscheiden von normalen Pflanzen.

## 77

Als weiteres visuelles Reprotergen wurde GFP, das grün-fluoreszierende Protein aus der Qualle *Aequorea victoria* eingesetzt. Dieses Reprotergen ist ein häufig verwendetes Reprotergen in Pflanzen (siehe z.B. Stewart, Plant Cell Rep 2001 20(5):376-82).

- 5 Ausgehend von *Arabidopsis thaliana* cDNA oder vom Plasmid pEGFP (BD Clontech, Heidelberg, Genbank-Eintrag U476561) wurde über PCR mittels der aufgeführten Oligonukleotide erzeugt. Dabei wurde folgendes Protokoll eingesetzt:

10 Zusammensetzung des PCR-Ansatzes (50 µL):

- 5,00 µL Template cDNA oder genomische DNA (ca. 1 µg)
- 5,00 µL 10x Puffer (Advantage-Polymerase) + 25 mM MgCl<sub>2</sub>
- 5,00 µL 2mM dNTP
- 15 1,25 µL je Primer (10 pmol/µL)
- 0,50 µL Advantage-Polymerase (Clontech)

PCR-Programm: Anfangsdenaturierung für 2 min bei 95°C, dann 35 Zyklen mit 45 sec 95°C, 45 sec 55°C und 2 min 72°C. Abschliessende

20 Extension von 5 min bei 72°C.

a) Ausgangsvektor pGEM-Toc159: Ausgehend von *Arabidopsis* cDNA wurde mit nachfolgendem Oligonukleotid-Primerpaar ein Fragment aus Toc159 (Genbank Acc.-No. **T14P8.24**) amplifiziert:

25

ONP18 (SEQ ID NO: 123):

5'-CTCGAGGAATTCATGGACTCAAAGTCGGTTACTCCA

ONP19 (SEQ ID NO: 124):

30 5'-GGATCCATAAGCAAGCTTTCTCACTCTCCCCATCTGTGGA

Das PCR Produkt wurde in den Vektor pGEM-T easy von Promega (Mannheim) gemäss Herstellerangaben kloniert, resultierend in dem pGEM-Toc159 Vektor und die Sequenz überprüft.

35

b) Ausgangsvektor pGEM-GFP: Ausgehend von dem Plasmid pEGFP (BD Clontech, Heidelberg, GenbankAcc.-No.: U476561) wurde mit nachfolgendem Oligonukleotid-Primerpaar ein Fragment aus GFP amplifiziert:

40

ONP20 (SEQ ID NO: 125): 5'-AAGCTTCCAACACTTGTCCTACTTTT

ONP21 (SEQ ID NO: 126): 5'-GGATCCTTAAAGCTCATCATGTTTGT

Das PCR Produkt wurde in den Vektor pGEM-T easy von Promega

- 45 (Mannheim) gemäss Herstellerangaben kloniert, resultierend in dem pGEM-GFP Vektor und die Sequenz überprüft.

c) Herstellung des Konstruktes pGEM-159-GFP Der Vektor pGEM-GFP wurde mit den Restriktionsenzymen HindIII und BamHI für 2 Stunden inkubiert. Parallel wurde der Vektor pGEM-Toc159 mit den gleichen Restriktionsenzymen inkubiert, anschliessend dann zusätzlich für 15 min mit alkalischer Phosphatase behandelt. Die alkalische Phosphatase wurde anschliessend durch Erhitzen auf 95 °C für 10 min inaktiviert. Die entstandenen DNA-Fragmente aus beiden Ansätzen wurden über Agarose-Gelelektrophorese aufgetrennt. Das 558 bp Fragment aus pGEM-GFP sowie das 3471 bp Fragment von pGEM-Toc159 wurden aus dem Gel ausgeschnitten und mit dem „Gelpurification“-Kit (Qiagen) nach Herstellerangaben aufgereinigt. Beide Fragmente wurden für 2 h bei 16°C ligiert (T4 Ligase, New England Biolabs) und anschliessend nach Herstellerangaben in E. coli DH5α Zellen (Stratagen) transformiert. Positive Klone wurden durch PCR mit dem Primerpaar OPN1 und OPN4 identifiziert und anschliessend verifiziert durch Sequenzierung. Der erhaltene Vektor wurde mit pGEM-159-GFP bezeichnet.

d) Herstellung des Suppressionskonstruktes 1: Der Vektor pGEM-159-GFP wurde einerseits mit den Restriktionsenzymen XhoI und BamHI, ein weiterer Ansatz mit BamHI und SalI inkubiert. Der zweite Ansatz mit BamHI/ SalI wurde anschliessend für weitere 15 min mit alkalischer Phosphatase inkubiert. Die DNA-Fragmente aus beiden Ansätzen wurden über Agarose-Gelelektrophorese aufgetrennt und folgende Fragmente ausgeschnitten: Ansatz BamHI-XhoI das 1091 bp Fragment; Ansatz BamHI-SalI das 4029 bp Fragment. Beide Fragmente wurden nach Aufreinigung aus dem Agarose-Gel (siehe oben) für 2 h bei 16°C mit T4 Ligase inkubiert und anschliessend in E. coli DH5α Zellen (Stratagen) transformiert. Positive Klone wurden durch PCR mit dem Primerpaar OPN1 identifiziert und anschliessend verifiziert durch Sequenzierung. Der erhaltene Vektor wurde als Suppressionskonstrukt 1 bezeichnet.

e) Herstellung des Suppressionskonstruktes 2: Das Suppressionskonstrukt 1 und der Vektor p3300.1 (Andreas Hilbrunner, Dissertation ETH Zürich, 2003) wurden für 2h Stunden mit dem Restriktionsenzym EcoRI inkubiert. Anschliessend wurde der Vektor p3300.1 15 min mit alkalischer Phosphatase behandelt. Beide Ansätze wurden gemischt und für 2 h bei 16°C mit T4 Ligase inkubiert. Der Ligationsansatz wurde dann in E. coli DH5α Zellen (Stratagen) transformiert. Das entstandene Suppressionskonstrukt 2 wurde dann für die Agrobacterium- und Pflanzentransformation eingesetzt. Die Nukleinsäuresequenz kodierend für Suppressionskonstrukt 2 (p3300.1-Toc159-GFP-RNAi) ist unter SEQ ID NO: 122 wiedergegeben.

- Die Transformation von Agrobakterien und Pflanzen wurde wie in Beispiel 3 bzw. 4 beschrieben durchgeführt. Zum Nachweis der Funktionalität des Suppressionskonstruktes 2 wurde dieses durch die nach Bechtold et al., 1993 (C.R. Acad. Sci. Ser. III Sci. Vie., 316, 1194-1199) beschriebene Blüten-Transformationsmethode in *Arabidopsis* transformiert. Aus Ausgangsmaterial wurden *Arabidopsis* Pflanzen der Varietät Columbia-0 verwendet, die bereits die T-DNA des binären Vektors pBIN-35S-GFP enthielten.
- 10 Durch Anregung durch ultraviolettes Licht im Wellenlängenbereich 470-490 nm die grüne Fluoreszenz von GFP in diesen Pflanzen angeregt werden und damit die Expression des eingebrachten Transgens überprüft werden. Dazu wurden Keimlinge 1 Woche nach Keimung oder Blattstücke bei älteren Pflanzen mit dem Fluoreszenzmikroskop
- 15 MZFLIII von Leica analysiert. Zur Anregung von GFP wurden folgende Parameter eingestellt: Quecksilberlampe HBO 100W/DC, Filter GFP3, Bildbearbeitung Leica-Software. Speziell die Verwendung eines Filters (GFP3), der oberhalb einer Wellenlänge von 525 nm nicht mehr durchlässig ist, ermöglicht die GFP-Analyse von grünen
- 20 Blattmaterial. Ohne diesen Filter könnte die starke Autofluoreszenz des Blattfarbstoffes Chlorophyll nicht ausgeschlossen werden. Die zur Transformation verwendete *Arabidopsis* Linie zeigte eine starke GFP Expression nach mikroskopischer Analyse.
- 25 Transformierte Samen wurden direkt auf Erde ausgelegt und angezogen. Nach einer Woche wurde nach Keimlingen gesucht, die keinen oder einen reduzierten Anteil des Blattfarbstoffes Chlorophyll enthielten. Solche Pflanzen waren leicht an ihrer hellgrünen oder weissen Erscheinungsbild zu erkennen. Diese Pflanzen wurden dann
- 30 weiter durch Fluoreszenz-Mikroskopie untersucht und mit entsprechend parallel gewachsenen grünen Pflanzen verglichen. Fig.5A zeigt beispielhaft eine solche identifizierte Pflanze, die sich deutlich in der Farbe der Blätter von parallel gewachsenen Pflanzen unterscheidet. Dabei ist der Albino-Phänotyp (weisse Blätter)
- 35 auf die Wirkung des Toc159-Suppressionskonstrukts zurückzuführen. Die nicht transformierten Nachkommen der mit *Agrobacterium*-Suspension behandelten Pflanzen zeigen den Albino-Phänotyp nicht. Der auftretende Albino-Phänotyp ist damit ein spezifischer Effekt des eingebrachten Suppressionskonstruktes.
- 40 Die Fluoreszenz-mikroskopische Untersuchung der Albino-Pflanzen zeigte dann (Fig.5B), dass keine GFP-Signale in solchen Pflanzen gefunden werden konnte. Im Vergleich dazu zeigten die parallel gewachsenen grünen Pflanzen deutliche GFP Signale. Die Abwesen-
- 45 heit des GFP-Signals in allen identifizierten Albino-Pflanzen demonstriert die Funktionalität des Suppressionskonstruktes, denn nur die mit dem Suppressionkonstrukt transformierten Pflanzen



80

zeigten keine GFP-Signale mehr. Es konnte keine Segregation der beiden angestrebten Phänotypen beobachtet werden. Damit konnte gezeigt werden, dass durch Verwendung von nur einem Kontrollelement (Promotor) zwei funktionell völlig unterschiedliche Gene, 5 die ihrerseits durch unterschiedliche Kontrollelemente in ihrer Expression reguliert werden, ausgeschaltet werden konnten.

10

15

20

25

30

35

40

45

## Patentansprüche

1. Verfahren zur Verminderung der Expression von mindestens zwei  
5 verschiedenen, endogenen Zielgenen in einer eukaryotischen  
Zelle oder einem eukaryotischen Organismus durch Einbringen  
eines zumindest teilweise doppelsträngigen Ribonukleinsäure-  
moleküls in besagte eukaryotische Zelle oder besagten euka-  
ryotischen Organismus, wobei das doppelsträngige Ribonuklein-  
10 säuremolekül umfasst
  - a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei je-  
weils mindestens eine dieser "sense"-Ribonukleotidsequen-  
zen im wesentlichen identisch ist zu mindestens einem  
15 Teil des "sense"-RNA-Transkriptes eines jeden der besag-  
ten endogenen Zielgene und
  - b) "antisense"-Ribonukleotidsequenzen, die zu besagten  
"sense"-Ribonukleotidsequenzen unter a) im wesentlichen  
20 komplementären sind.
2. Verfahren nach Anspruch 1, wobei die transkribierten RNAs von  
mindestens zwei der in ihrer Expression verminderten Zielgene  
untereinander eine Homologie von unter 90% haben.  
25
3. Verfahren nach Anspruch 1 oder 2, wobei die doppelsträngige  
RNA durch ein einziges selbstkomplementäres Ribonukleotidmo-  
lekül gebildet wird.
- 30 4. Verfahren nach einem der Ansprüche 1 bis 3, wobei mindestens  
eine der ausgehend von den einzelnen "sense"-Ribonukleotidse-  
quenzen gebildeten doppelsträngigen RNA-Strukturen eine Länge  
eines geradzahligen Vielfachen von 21 oder 22 Basenpaaren  
hat.  
35
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei das Ribonu-  
kleotidmolekül zwischen mindestens einer "sense"-Ribonukleo-  
tidsequenz und der dazu im wesentlichen komplementären "anti-  
sense"-Ribonukleotidsequenz eine Ribonukleotidsequenz kodie-  
40 rend für ein Intron enthält.

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass mindestens zwei der endogenen Zielgene ausgewählt sind aus jeweils unterschiedlichen Klassen von Speicherprotein ausgewählt aus den Speicherprotein-Klassen der
- 5 2S-Albumine, 7S-Globuline, 11S/12S-Globuline oder Zein-Proline.
7. Verfahren nach einem der Ansprüche 1 bis 6, wobei mindestens eine "sense"-Ribonukleotidsequenz im wesentlichen identisch
- 10 ist zu mindestens einem Teil des "sense"-RNA-Transkriptes
- a) einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 59, 61,
- 15 63, 65, 67, 69, 71, 93, 95, 97, 99, 101, 103, 105, 107, 109 oder 112, oder
- b) eines Gens aus dem Homogentisatabbauweg gemäß SEQ ID NO: 115, 116, 118 oder 120, oder
- 20 c) eines Gens ausgewählt aus der Gruppe bestehend aus Acetyltransacylasen, Acyltransportproteinen, Fettsäuredesaturasen, Malonyltransacylasen,  $\beta$ -Ketoacyl-ACP-synthetasen, 3-Keto-ACP-reduktasen, Enoyl-ACP-hydrasen, Thioesterasen, Enoyl-ACP-reduktasen, ADP-Glucosepyrophosphorylasen, Phosphorylasen, Stärkesynthetasen, Q-Enzymen, Sucrose-6-phosphatsynthetasen, Sucrose-6-phosphatphosphatasen, ADP-Glucosepyrophosphorylasen, Branching-Enzymen, Debranching-Enzymen, Amylasen, Chalconsynthasen, Chalconisomerasen, Phenylalaninammonialyasen, Dehydrokaempferol(flavone)hydroxylasen, Dihydroflavonolreduktasen, Dihydroflavanol-2-hydroxylasen, Flavonoid-3'-hydroxylasen, Flavonoid-5'-hydroxylasen, Flavonoidglycosyltransferasen, Flavonoidmethyltransferasen, Flavonoidacyltransferasen,
- 25 Polygalacturonasen, Cellulasen, Pectinesterasen,  $\beta$ -(1-4)Glucanasen,  $\beta$ -Galactanasen, 1-Aminocyclopropan-1-carboxylatsynthasen, Phytoendesaturasen, Cinnamoyl-CoA:NADPH-Reduktasen, Cinnamoylalkoholdehydrogenasen, Coffeinsäure-O-methyltransferasen, Cinnamoylalkoholdehydrogenasen, Polyphenoloxidasen, Homogentisat-1,2-dioxygenasen, Maleylacetoacetatisomerasen, Fumarylacetoacetylhydrolasen, N-Methyl-putrescinoxidasen, Putrescin-N-methyltransferasen, 7-Methylxanthine-3-methyltransferasen, 1-Methylxanthin-3-methyltransferasen und Threoninsynthasen.
- 30 35 40 45

8. Ribonukleinsäuremolekül, das eine zumindest teilweise doppelsträngige Struktur hat und umfasst
- 5 a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jeweils mindestens eine dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens, wobei jedoch nicht alle "sense"-Ribonukleotidsequenzen zu dem "sense"-RNA-Transkript eines einzigen endogenen Zielgens im wesentlichen identisch sind, und
- 10 b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.
- 15 9. Ribonukleinsäuremolekül nach Anspruch 8, wobei das Ribonukleinsäuremolekül wie in einem der Ansprüche 2 bis 7 gekennzeichnet ist.
- 20 10. Transgene Expressionskassette enthaltend in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für doppelsträngiges Ribonukleinsäuremolekül gemäß einem der Ansprüche 8 oder 9, wobei das Ribonukleinsäuremolekül aus einem einzigen RNA-Strang gebildet wird.
- 25 11. Transgenes Expressionssystem enthaltend
- 30 a) in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für "sense"-Ribonukleotidsequenzen eines doppelsträngigen Ribonukleinsäuremoleküls gemäß einem der Ansprüche 8 oder 9 und
- 35 b) in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für "antisense"-Ribonukleotidsequenzen eines doppelsträngigen Ribonukleinsäuremoleküls gemäß einem der Ansprüche 8 oder 9,
- 40 wobei das Ribonukleinsäuremolekül aus den beiden unter a) und b) definierten Strängen gebildet wird, und die Promotoren so gewählt sind, das in einem bestimmten Organismus oder Zelle die gleichzeitige Expression von "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen gewährleistet ist.
- 45

## 84

12. Transgener Vektor enthaltend eine transgene Expressionskassette gemäß Anspruch 10 oder ein transgenes Expressionssystem gemäß Anspruch 11.
- 5 13. Transgener Organismus enthaltend eine transgene Expressionskassette gemäß Anspruch 10 oder ein transgenes Expressionssystem gemäß Anspruch 11 oder einen transgenen Vektor gemäß Anspruch 12.
- 10 14. Transgener Organismus nach Anspruch 13 ausgewählt aus der Gruppe bestehend aus Bakterien, Hefen, nicht-menschlichen Tieren und Pflanzen.
- 15 15. Transgener Organismus nach Ansprüchen 13, dadurch gekennzeichnet, dass die Pflanze ausgewählt ist aus der Gruppe der landwirtschaftlichen Nutzpflanzen.
- 20 16. Verwendung eines Ribonukleotidmoleküls nach einem der Ansprüche 8 oder 9, einer transgenen Expressionskassette gemäß Anspruch 10, eines transgenen Expressionssystems gemäß Anspruch 11, eines transgenen Vektors gemäß Anspruch 12 oder eines transgenen Organismus gemäß einem der Ansprüche 13 bis 15 zur Herstellung von Arzneimitteln, in biotechnologischen Verfahren oder in der Pflanzenbiotechnologie.
- 25 17. Verwendung nach Anspruch 16, wobei mindestens einer der nachfolgenden Eigenschaften in Pflanzen erzielt wird:
- 30 a) Verbesserter Schutz gegen abiotische Stressfaktoren
- b) Modifikation der Zusammensetzung und/oder des Gehaltes an Fettsäuren, Lipiden oder Ölen
- 35 c) Modifikation der Kohlenhydratzusammensetzung
- d) Veränderung der Farbe oder Pigmentierung
- e) Verminderung des Gehaltes von Speicherproteinen
- 40 f) Erreichen einer Resistenz gegen pflanzliche Pathogene
- g) Verhinderung von Halmbruch
- h) Verzögerung der Fruchtreifung
- 45 i) Erzielen einer männlichen Sterilität

## 85

- j) Verminderung unerwünschter oder toxischer Pflanzeninhaltsstoffe
- 5 k) Verzögerung von Alterserscheinungen
- l) Modifikation der Lignifikation und/oder des Ligningehaltes
- 10 m) Modifikation des Faseranteils in Nahrungsmitteln oder der Faserqualität in Baumwolle
- n) Verminderung der Stoßanfälligkeit
- 15 o) Steigerung der Vitamin E Biosynthese
- p) Verminderung des Nikotingehaltes, des Coffeingehaltes oder des Theophyllin-Gehaltes
- 20 q) Erhöhung des Methioningehaltes durch Verminderung der Threoninbiosynthese

25

30

35

40

45

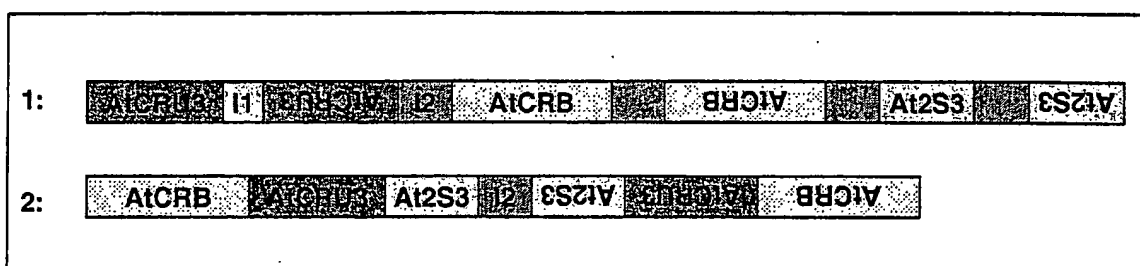


Fig. 1

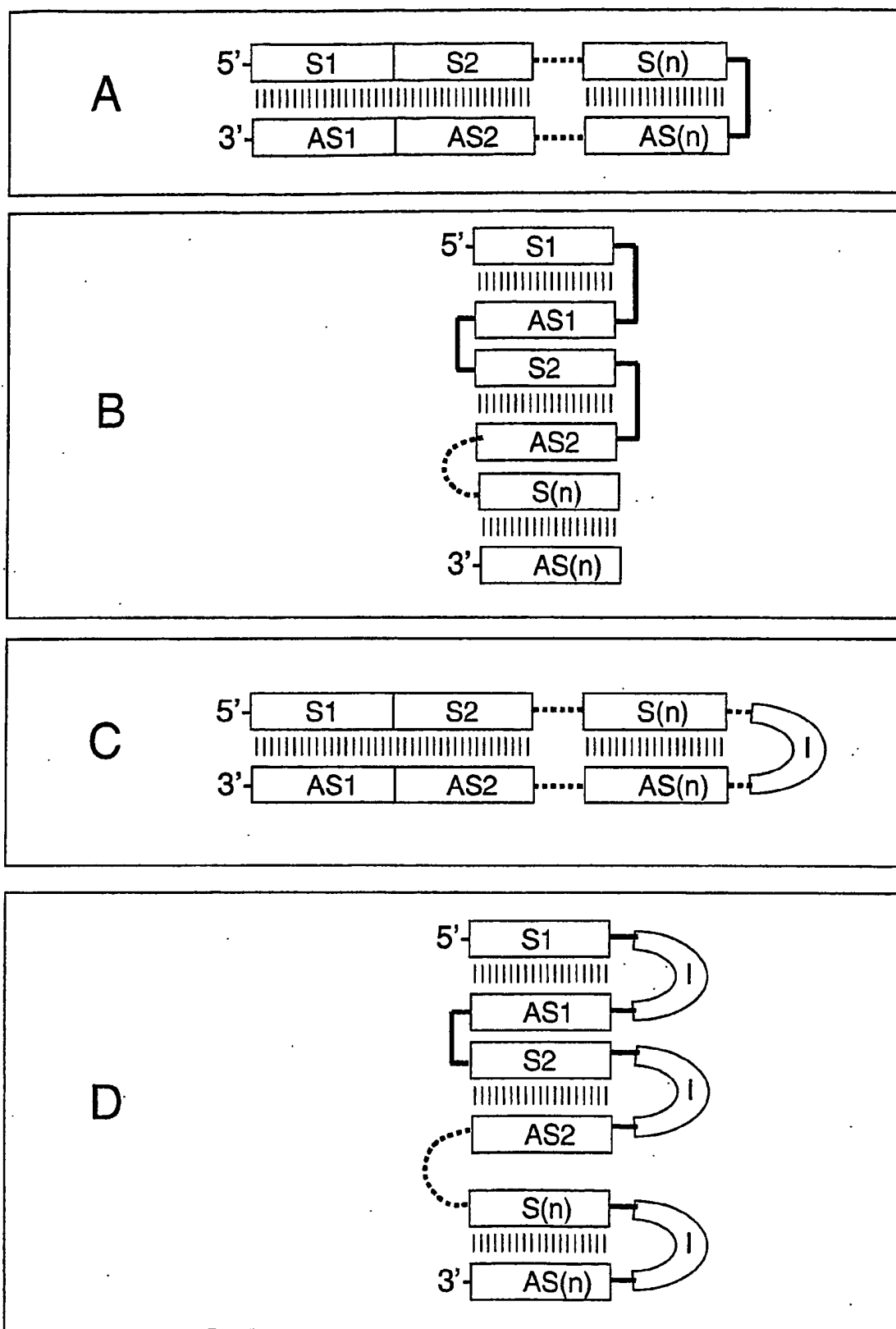


Fig.2



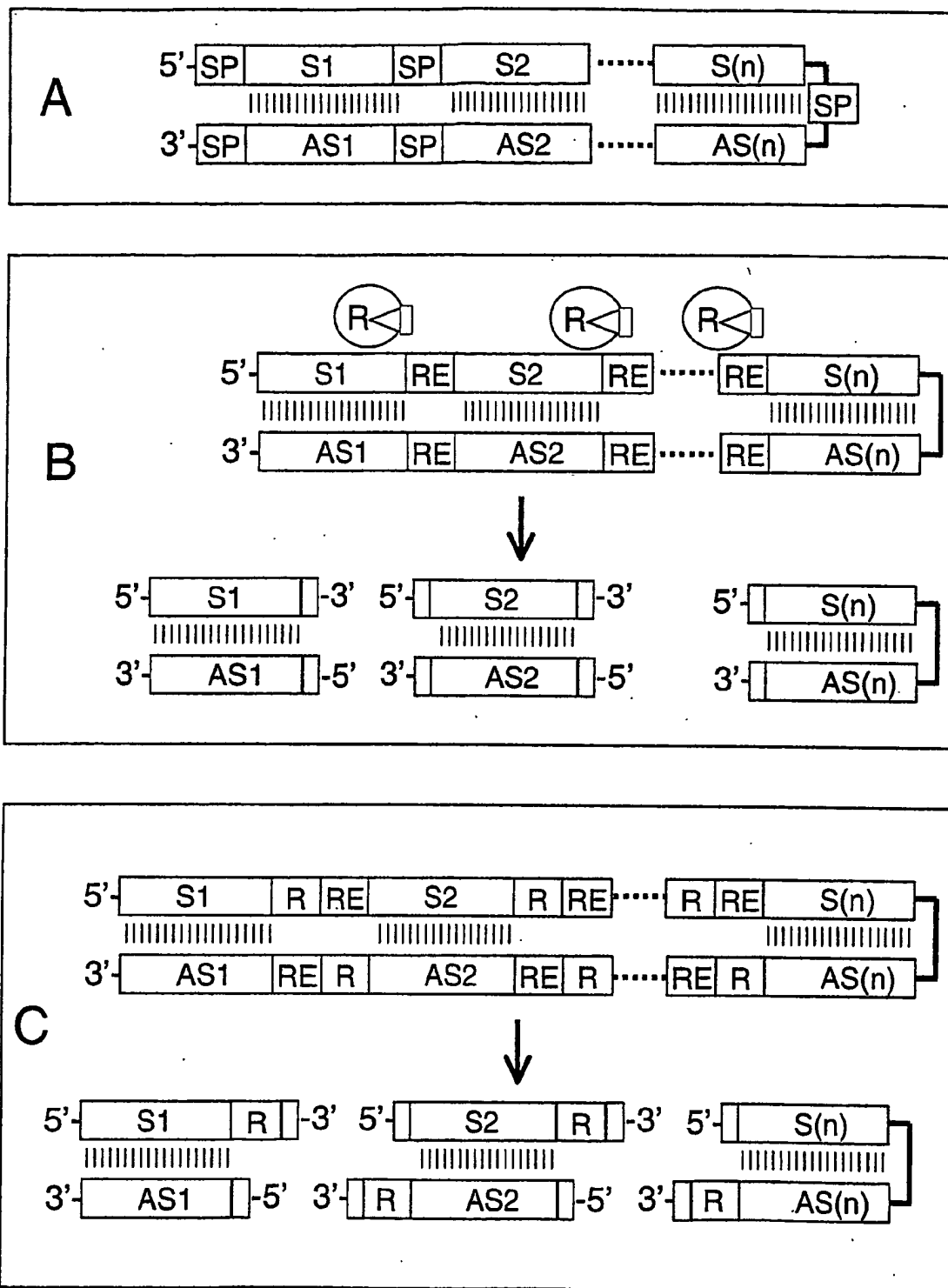


Fig.3

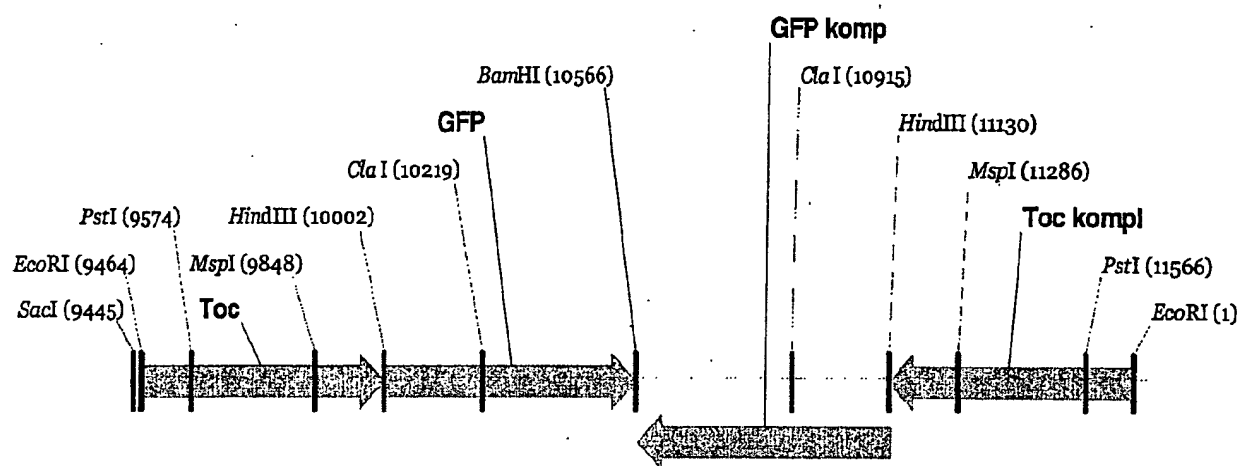
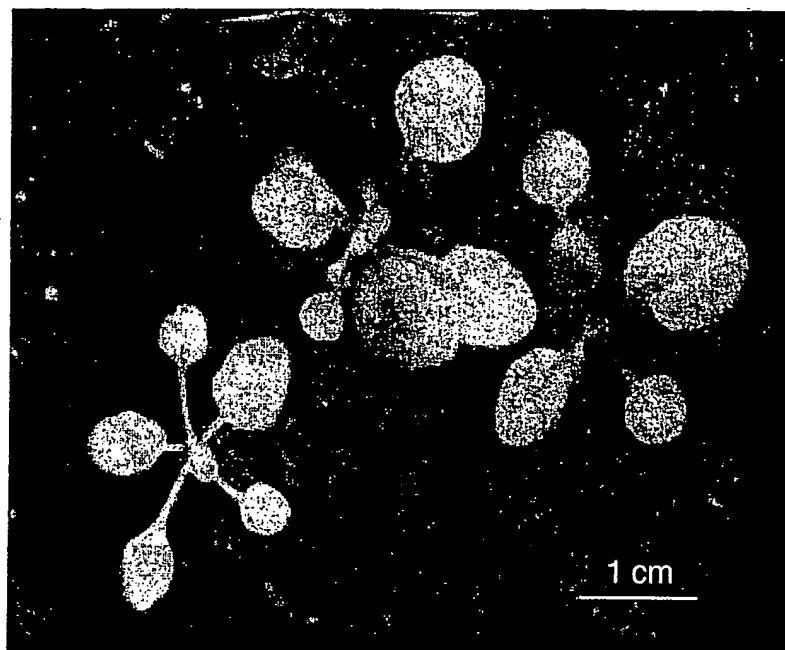


Fig.4

A



B

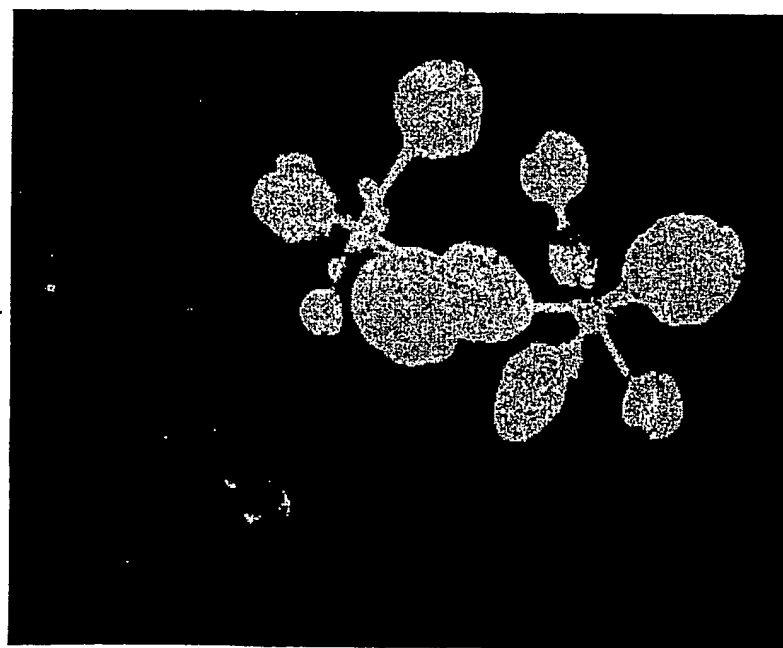


Fig.5

## SEQUENZPROTOKOLL

&lt;110&gt; BASF Plant Science GmbH

&lt;120&gt; Konstrukte und Verfahren zur Regulation der Genexpression

&lt;130&gt; PD009300062-AT

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 126

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 495

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(492)

&lt;223&gt; albumine 2S subunit 1

&lt;400&gt; 1

atg gca aac aag ttg ttc ctc gtc tgc gca gct ctc gct ctc tgc ttc	48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe	
1 5 10 15	
ctc ctc acc aac gct tcc atc tac cgc acc gtc gtt gag ttc gaa gaa	96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu	
20 25 30	
gat gac gcc act aac ccc ata ggc cca aaa atg agg aaa tgc cgc aag	144
Asp Asp Ala Thr Asn Pro Ile Gly Pro Lys Met Arg Lys Cys Arg Lys	
35 40 45	
gag ttt cag aaa gaa caa cac cta aga gct tgc cag caa ttg atg ctc	192
Glu Phe Gln Lys Glu Gln His Leu Arg Ala Cys Gln Gln Leu Met Leu	
50 55 60	
cag caa gca agg caa ggc cgt agc gat gag ttt gat ttc gaa gac gac	240
Gln Gln Ala Arg Gln Gly Arg Ser Asp Glu Phe Asp Phe Glu Asp Asp	
65 70 75 80	
atg gag aac cca cag gga caa cag cag gaa caa cag cta ttc cag cag	288
Met Glu Asn Pro Gln Gly Gln Gln Gln Glu Gln Gln Leu Phe Gln Gln	
85 90 95	
tgc tgc aac gag ctt cgc cag gaa gag cca gat tgt gtt tgc ccc acc	336
Cys Cys Asn Glu Leu Arg Gln Glu Glu Pro Asp Cys Val Cys Pro Thr	
100 105 110	
ttg aaa caa gct gcc aag gcc gtt aga ctc cag gga cag cac caa cca	384
Leu Lys Gln Ala Ala Lys Ala Val Arg Leu Gln Gly Gln His Gln Pro	
115 120 125	
atg caa gtc agg aaa att tac cag aca gcc aag cac ttg ccc aac gtt	432
Met Gln Val Arg Lys Ile Tyr Gln Thr Ala Lys His Leu Pro Asn Val	
130 135 140	
tgc gac atc ccg caa gtt gat gtt tgt ccc ttc aac atc cct tca ttc	480
Cys Asp Ile Pro Gln Val Asp Val Cys Pro Phe Asn Ile Pro Ser Phe	
145 150 155 160	
cct tct ttc tac taa	495
Pro Ser Phe Tyr	

<210> 2  
 <211> 164  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 2  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe  
 1 5 10 15  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu  
 20 25 30  
 Asp Asp Ala Thr Asn Pro Ile Gly Pro Lys Met Arg Lys Cys Arg Lys  
 35 40 45  
 Glu Phe Gln Lys Glu Gln His Leu Arg Ala Cys Gln Gln Leu Met Leu  
 50 55 60  
 Gln Gln Ala Arg Gln Gly Arg Ser Asp Glu Phe Asp Phe Glu Asp Asp  
 65 70 75 80  
 Met Glu Asn Pro Gln Gly Gln Gln Gln Glu Gln Gln Leu Phe Gln Gln  
 85 90 95  
 Cys Cys Asn Glu Leu Arg Gln Glu Glu Pro Asp Cys Val Cys Pro Thr  
 100 105 110  
 Leu Lys Gln Ala Ala Lys Ala Val Arg Leu Gln Gly Gln His Gln Pro  
 115 120 125  
 Met Gln Val Arg Lys Ile Tyr Gln Thr Ala Lys His Leu Pro Asn Val  
 130 135 140  
 Cys Asp Ile Pro Gln Val Asp Val Cys Pro Phe Asn Ile Pro Ser Phe  
 145 150 155 160  
 Pro Ser Phe Tyr

<210> 3  
 <211> 495  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (1)..(492)  
 <223> albumine 2S subunit 3  
 <400> 3  
 atg gct aac aag ctc ttc ctc gtc tgc gca act ctc gcc ctc tgc ttc 48  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Leu Ala Leu Cys Phe  
 1 5 10 15  
 ctc ctc acc aac gct tcc atc tac cgc acc gtt gtc gaa ttc gaa gaa 96  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu  
 20 25 30  
 gat gac gcc agc aac ccc gta ggt cca aga cag aga tgc cag aag gag 144  
 Asp Asp Ala Ser Asn Pro Val Gly Pro Arg Gln Arg Cys Gln Lys Glu  
 35 40 45  
 ttt cag caa tca caa cac cta aga gct tgc cag aga tgg atg agc aag 192  
 Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Arg Trp Met Ser Lys  
 50 55 60

```

caa atg agg caa gga cgt ggt ggt ggt cct tcc ctc gac gat gag ttc 240
Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
 65          70          75          80
gat ttg gag ggc ccc cag cag gga tac cag cta ctc cag cag tgc tgc 288
Asp Phe Glu Gly Pro Gln Gln Gly Tyr Gln Leu Leu Gln Gln Cys Cys
          85          90          95
aac gag ctt cgc cag gaa gag cca gtt tgc gtt tgc ccc acc ttg aaa 336
Asn Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro Thr Leu Lys
          100          105          110
caa gct gcc agg gca gtt agc ctc cag gga cag cac gga cca ttc caa 384
Gln Ala Ala Arg Ala Val Ser Leu Gln Gly Gln His Gly Pro Phe Gln
          115          120          125
tcc agg aaa att tac cag tca gct aag tac ttg cct aac att tgc aag 432
Ser Arg Lys Ile Tyr Gln Ser Ala Lys Tyr Leu Pro Asn Ile Cys Lys
          130          135          140
atc cag caa gtt ggt gaa tgt ccc ttc cag acc acc atc cct ttc ttc 480
Ile Gln Gln Val Gly Glu Cys Pro Phe Gln Thr Thr Ile Pro Phe Phe
          145          150          155          160
cct cct tac tac tag
Pro Pro Tyr Tyr 495

<210> 4
<211> 164
<212> PRT
<213> Arabidopsis thaliana

<400> 4
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Leu Ala Leu Cys Phe
 1          5          10          15
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu
          20          25          30
Asp Asp Ala Ser Asn Pro Val Gly Pro Arg Gln Arg Cys Gln Lys Glu
          35          40          45
Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Arg Trp Met Ser Lys
          50          55          60
Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
          65          70          75          80
Asp Phe Glu Gly Pro Gln Gln Gly Tyr Gln Leu Leu Gln Gln Cys Cys
          85          90          95
Asn Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro Thr Leu Lys
          100          105          110
Gln Ala Ala Arg Ala Val Ser Leu Gln Gly Gln His Gly Pro Phe Gln
          115          120          125
Ser Arg Lys Ile Tyr Gln Ser Ala Lys Tyr Leu Pro Asn Ile Cys Lys
          130          135          140
Ile Gln Gln Val Gly Glu Cys Pro Phe Gln Thr Thr Ile Pro Phe Phe
          145          150          155          160
Pro Pro Tyr Tyr

```

&lt;210&gt; 5

&lt;211&gt; 513

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(510)

&lt;223&gt; albumine 2S subunit 2

&lt;400&gt; 5

```

atg gca aac aag ctc ttc ctc gtc tgc gca act ttc gcc ctc tgc ttc   48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Phe Ala Leu Cys Phe
  1             5             10             15

ctc ctc acc aac gct tcc atc tac cgc act gtt gtc gag ttc gac gaa   96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
             20             25             30

gat gac gcc agc aac ccc atg ggc cca aga cag aaa tgt cag aag gag   144
Asp Asp Ala Ser Asn Pro Met Gly Pro Arg Gln Lys Cys Gln Lys Glu
             35             40             45

ttt cag caa tca cag cac cta aga gct tgc cag aaa ttg atg cgc atg   192
Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Lys Leu Met Arg Met
             50             55             60

caa atg agg caa ggc cgt ggt ggt ggt ccc tcc ctc gac gat gag ttc   240
Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
             65             70             75             80

gat ttg gaa gac gac atc gag aac cca caa ggc ccc cag cag gga cac   288
Asp Leu Glu Asp Asp Ile Glu Asn Pro Gln Gly Pro Gln Gln Gly His
             85             90             95

cag atc ctc cag cag tgc tgc agc gag ctt cgc cag gaa gag cca gtt   336
Gln Ile Leu Gln Gln Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val
             100            105            110

tgt gtt tgc ccc acc ttg aga caa gct gcc agg gcc gtt agc ctc cag   384
Cys Val Cys Pro Thr Leu Arg Gln Ala Ala Arg Ala Val Ser Leu Gln
             115            120            125

gga caa cac gga cca ttc caa tcc agg aaa att tac aag aca gct aag   432
Gly Gln His Gly Pro Phe Gln Ser Arg Lys Ile Tyr Lys Thr Ala Lys
             130            135            140

tac ttg cct aac att tgc aag atc cag caa gtt ggt gaa tgc ccc ttc   480
Tyr Leu Pro Asn Ile Cys Lys Ile Gln Gln Val Gly Glu Cys Pro Phe
             145            150            155            160

cag acc acc atc cct ttc ttc cct cct tac taa   513
Gln Thr Thr Ile Pro Phe Phe Pro Pro Tyr
             165            170

```

&lt;210&gt; 6

&lt;211&gt; 170

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 6

```

Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Phe Ala Leu Cys Phe
  1             5             10             15

Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
             20             25             30

Asp Asp Ala Ser Asn Pro Met Gly Pro Arg Gln Lys Cys Gln Lys Glu
             35             40             45

```

Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Lys Leu Met Arg Met  
 50 55 60  
 Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 Asp Leu Glu Asp Asp Ile Glu Asn Pro Gln Gly Pro Gln Gln Gly His  
 85 90 95  
 Gln Ile Leu Gln Gln Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val  
 100 105 110  
 Cys Val Cys Pro Thr Leu Arg Gln Ala Ala Arg Ala Val Ser Leu Gln  
 115 120 125  
 Gly Gln His Gly Pro Phe Gln Ser Arg Lys Ile Tyr Lys Thr Ala Lys  
 130 135 140  
 Tyr Leu Pro Asn Ile Cys Lys Ile Gln Gln Val Gly Glu Cys Pro Phe  
 145 150 155 160  
 Gln Thr Thr Ile Pro Phe Phe Pro Pro Tyr  
 165 170

<210> 7  
 <211> 501  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (1)..(498)  
 <223> albumine 2S subunit 4  
 <400> 7

atg gcg aac aag ctc ttc ctc gtc tgc gca gct ctc gcc ctg tgt ttc 48  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe  
 1 5 10 15  
 atc ctc acc aac gct tcc gtc tat cgc acc gtt gtc gag ttc gac gaa 96  
 Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu  
 20 25 30  
 gat gac gcc agt aac ccc ata ggc cca ata cag aaa tgt cag aag gag 144  
 Asp Asp Ala Ser Asn Pro Ile Gly Pro Ile Gln Lys Cys Gln Lys Glu  
 35 40 45  
 ttt cag caa gac cag cac cta aga gct tgc cag aga tgg atg cgc aag 192  
 Phe Gln Gln Asp Gln His Leu Arg Ala Cys Gln Arg Trp Met Arg Lys  
 50 55 60  
 caa atg tgg caa gga cgt ggt ggt ggt cct tcc ctc gac gat gag ttc 240  
 Gln Met Trp Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 gat atg gaa gac gac atc gag aac ccg cag aga cga cag cta ctc cag 288  
 Asp Met Glu Asp Asp Ile Glu Asn Pro Gln Arg Arg Gln Leu Leu Gln  
 85 90 95  
 aag tgc tgc agc gag ctt cgc caa gaa gag cca gtt tgc gtt tgc ccc 336  
 Lys Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro  
 100 105 110  
 acc ttg aga caa gct gcc aag gcc gtt aga ttc cag gga cag caa cac 384  
 Thr Leu Arg Gln Ala Ala Lys Ala Val Arg Phe Gln Gly Gln Gln His  
 115 120 125



caa cca gag caa gtc agg aaa att tac cag gca gct aag tac ttg cct 432  
 Gln Pro Glu Gln Val Arg Lys Ile Tyr Gln Ala Ala Lys Tyr Leu Pro  
 130 135 140

aac att tgc aaa atc cag caa gtt ggt gtt tgc ccc ttc cag atc cct 480  
 Asn Ile Cys Lys Ile Gln Gln Val Gly Val Cys Pro Phe Gln Ile Pro  
 145 150 155 160

tca atc cct tct tac tac taa 501  
 Ser Ile Pro Ser Tyr Tyr  
 165

<210> 8  
 <211> 166  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 8  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe  
 1 5 10 15  
 Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu  
 20 25 30  
 Asp Asp Ala Ser Asn Pro Ile Gly Pro Ile Gln Lys Cys Gln Lys Glu  
 35 40 45  
 Phe Gln Gln Asp Gln His Leu Arg Ala Cys Gln Arg Trp Met Arg Lys  
 50 55 60  
 Gln Met Trp Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 Asp Met Glu Asp Asp Ile Glu Asn Pro Gln Arg Arg Gln Leu Leu Gln  
 85 90 95  
 Lys Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro  
 100 105 110  
 Thr Leu Arg Gln Ala Ala Lys Ala Val Arg Phe Gln Gly Gln Gln His  
 115 120 125  
 Gln Pro Glu Gln Val Arg Lys Ile Tyr Gln Ala Ala Lys Tyr Leu Pro  
 130 135 140  
 Asn Ile Cys Lys Ile Gln Gln Val Gly Val Cys Pro Phe Gln Ile Pro  
 145 150 155 160  
 Ser Ile Pro Ser Tyr Tyr  
 165

<210> 9  
 <211> 1473  
 <212> DNA  
 <213> Brassica napus

<220>  
 <221> CDS  
 <222> (1)..(1470)  
 <223> cruciferin

<400> 9  
 atg gct cgg ctc tca tct ctt ctc tct ttt tcc tta gca ctt ttg atc 48  
 Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Ile  
 1 5 10 15

ttt ctc cat ggc tct aca gct caa cag ttt cca aac gag tgt cag cta	96
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu	
20 25 30	
gac cag ctc aat gca ctg gag ccg tca cac gta ctt aag gct gag gct	144
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala	
35 40 45	
ggc cgc atc gag gtg tgg gac cac cac gct cct cag cta cgt tgc tct	192
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	
ggc gtc tcc ttt gta cgt tac atc atc gag tct aag ggt ctc tac ttg	240
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu	
65 70 75 80	
ccc tct ttc ttt agc acc gcg aag ctc tcc ttc gtg gct aaa gga gaa	288
Pro Ser Phe Phe Ser Thr Ala Lys Leu Ser Phe Val Ala Lys Gly Glu	
85 90 95	
ggc ctt atg ggg aga gtg gtc cct gga tgc gcc gag aca ttc cag gac	336
Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp	
100 105 110	
tca tca gtg ttt caa cca agc ggt ggt agc ccc tcg gga gaa ggt cag	384
Ser Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Ser Gly Glu Gly Gln	
115 120 125	
ggc caa gga caa caa ggt cag ggc caa ggc cac caa ggt caa ggc caa	432
Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln	
130 135 140	
gga caa cag ggc caa caa ggt cag caa gga caa cag agt caa ggc cag	480
Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln	
145 150 155 160	
ggc ttc cgt gat atg cac cag aaa gtg gag cac ata agg act ggg gac	528
Gly Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp	
165 170 175	
acc atc gct aca cat ccc ggt gta gcc caa tgg ttc tac aac gac gga	576
Thr Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly	
180 185 190	
aac caa cca ctt gtc atc gtt tcc gtc ctc gat tta gcc agc cac cag	624
Asn Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln	
195 200 205	
aat cag ctc gac cgc aac cca agg cca ttt tac tta gcc gga aac aac	672
Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn	
210 215 220	
cca caa ggc caa gta tgg ata gaa gga cgc gag caa cag cca caa aag	720
Pro Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys	
225 230 235 240	
aac atc ctt aat ggc ttc aca cca gag gtt ctt gct aaa gct ttc aag	768
Asn Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys	
245 250 255	
atc gat gtt agg aca gcg caa caa ctt cag aac cag caa gac aac cgt	816
Ile Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg	
260 265 270	
gga aac att atc cga gtc caa ggc cca ttc agt gtc att agg ccg cct	864
Gly Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro	
275 280 285	

```

ttg agg agt cag aga ccg cag gag aca gaa gtt aac ggt tta gaa gag 912
Leu Arg Ser Gln Arg Pro Gln Glu Thr Glu Val Asn Gly Leu Glu Glu
290 295 300

acc ata tgc agc gcg agg tgc acc gat aac ctc gat gac cca tct aat 960
Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn
305 310 315 320

gct gac gta tac aag cca cag ctc ggt tac atc agc act ctg aac agc 1008
Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser
325 330 335

tat gat ctc ccc atc ctt cgc ttc ctt cgt ctc tca gcc ctc cgt gga 1056
Tyr Asp Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly
340 345 350

tct atc cgt caa aac gcg atg gtg ctt cca cag tgg aac gca aac gca 1104
Ser Ile Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala
355 360 365

aac gcg gtt ctc tac gtg aca gac ggg gaa gcc cat gtg cag gtg gtt 1152
Asn Ala Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val
370 375 380

aac gac aac ggt gac aga gtg ttc gac gga caa gtc tct caa gga cag 1200
Asn Asp Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln
385 390 395 400

cta ctt tcc ata cca caa ggt ttc tcc gtg gtg aaa cgc gca aca agc 1248
Leu Leu Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser
405 410 415

gaa cag ttc cgg tgg atc gag ttc aag aca aac gca aac gca cag atc 1296
Glu Gln Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile
420 425 430

aac aca ctt gct gga cga acc tcg gtc ttg aga ggt tta cca tta gag 1344
Asn Thr Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu
435 440 445

gtc ata tcc aat ggg tac caa atc tca ctc gaa gaa gca aga agg gtt 1392
Val Ile Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val
450 455 460

aag ttc aac acg atc gag acc act ttg acg cac agc agt ggc cca gct 1440
Lys Phe Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala
465 470 475 480

agc tac gga ggg cca agg aag gct gat gct taa 1473
Ser Tyr Gly Gly Pro Arg Lys Ala Asp Ala
485 490

```

&lt;210&gt; 10

&lt;211&gt; 490

&lt;212&gt; PRT

&lt;213&gt; Brassica napus

&lt;400&gt; 10

```

Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Ile
1 5 10 15

```

```

Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu
20 25 30

```

```

Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala
35 40 45

```

Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser  
 50 55 60  
 Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu  
 65 70 75 80  
 Pro Ser Phe Phe Ser Thr Ala Lys Leu Ser Phe Val Ala Lys Gly Glu  
 85 90 95  
 Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp  
 100 105 110  
 Ser Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Ser Gly Glu Gly Gln  
 115 120 125  
 Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln  
 130 135 140  
 Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln  
 145 150 155 160  
 Gly Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp  
 165 170 175  
 Thr Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly  
 180 185 190  
 Asn Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln  
 195 200 205  
 Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn  
 210 215 220  
 Pro Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys  
 225 230 235 240  
 Asn Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys  
 245 250 255  
 Ile Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg  
 260 265 270  
 Gly Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro  
 275 280 285  
 Leu Arg Ser Gln Arg Pro Gln Glu Thr Glu Val Asn Gly Leu Glu Glu  
 290 295 300  
 Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn  
 305 310 315 320  
 Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser  
 325 330 335  
 Tyr Asp Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly  
 340 345 350  
 Ser Ile Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala  
 355 360 365  
 Asn Ala Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val  
 370 375 380  
 Asn Asp Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln  
 385 390 395 400  
 Leu Leu Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser  
 405 410 415  
 Glu Gln Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile  
 420 425 430

## 110

&lt;222&gt; (1)..(642)

&lt;223&gt; avenin

&lt;400&gt; 107

```

atg aag atc ttc ttc ttc tta gct ctc ctt gct ctg gta gtg agc gcc 48
Met Lys Ile Phe Phe Phe Leu Ala Leu Leu Ala Leu Val Val Ser Ala
  1             5             10             15

acc ttt gca caa tat gca gaa tct gac ggt agt tat gag gaa gtg gag 96
Thr Phe Ala Gln Tyr Ala Glu Ser Asp Gly Ser Tyr Glu Glu Val Glu
             20             25             30

ggt tct cat gat cga tgc caa caa cat cag atg aag ctg gac tct tgc 144
Gly Ser His Asp Arg Cys Gln Gln His Gln Met Lys Leu Asp Ser Cys
             35             40             45

aga gag tac gtg gcg gag cgg tgc aca acg atg aga gat ttt ccg atc 192
Arg Glu Tyr Val Ala Glu Arg Cys Thr Thr Met Arg Asp Phe Pro Ile
             50             55             60

acc tgg cca tgg aaa tgg tgg aag ggt ggt tgc gag gag ctc cgc aat 240
Thr Trp Pro Trp Lys Trp Trp Lys Gly Gly Cys Glu Glu Leu Arg Asn
             65             70             75             80

gag tgc tgc caa ctg ttg ggc cag atg cca tcg gag tgt cgc tgt gat 288
Glu Cys Cys Gln Leu Leu Gly Gln Met Pro Ser Glu Cys Arg Cys Asp
             85             90             95

gcg att tgg aga tca atc cag cgc gag ctt ggt ggc ttc ttt gga act 336
Ala Ile Trp Arg Ser Ile Gln Arg Glu Leu Gly Gly Phe Phe Gly Thr
             100             105             110

caa caa ggt ctg ata ggg aaa agg ttg aag ata gcc aag agt ttg ccc 384
Gln Gln Gly Leu Ile Gly Lys Arg Leu Lys Ile Ala Lys Ser Leu Pro
             115             120             125

acg cag tca aca tgg gcc ctg agt gca ata tcc cca aac tcc atg gtt 432
Thr Gln Ser Thr Trp Ala Leu Ser Ala Ile Ser Pro Asn Ser Met Val
             130             135             140

agc cac att gct gga aag agc tcc att ctt cgt gcc ttg ccc gtg gat 480
Ser His Ile Ala Gly Lys Ser Ser Ile Leu Arg Ala Leu Pro Val Asp
             145             150             155             160

gtc ctc gcc aat gca tac cgc att tcc agg caa gaa gcc cga aac ctc 528
Val Leu Ala Asn Ala Tyr Arg Ile Ser Arg Gln Glu Ala Arg Asn Leu
             165             170             175

aaa aac aac agg gga caa gag tct ggt gta ttc act cca aaa ttt acc 576
Lys Asn Asn Arg Gly Gln Glu Ser Gly Val Phe Thr Pro Lys Phe Thr
             180             185             190

caa acg agc ttc caa cct tat cca gag ggc gag gat gag tca tct ttg 624
Gln Thr Ser Phe Gln Pro Tyr Pro Glu Gly Glu Asp Glu Ser Ser Leu
             195             200             205

att aat aag gca tca gag taa 645
Ile Asn Lys Ala Ser Glu
             210

```

&lt;210&gt; 108

&lt;211&gt; 214

&lt;212&gt; PRT

&lt;213&gt; Avena sativa

&lt;400&gt; 108

```

Met Lys Ile Phe Phe Phe Leu Ala Leu Leu Ala Leu Val Val Ser Ala
  1             5             10             15

```

## 111

Thr Phe Ala Gln Tyr Ala Glu Ser Asp Gly Ser Tyr Glu Glu Val Glu  
 20 25 30  
 Gly Ser His Asp Arg Cys Gln Gln His Gln Met Lys Leu Asp Ser Cys  
 35 40 45  
 Arg Glu Tyr Val Ala Glu Arg Cys Thr Thr Met Arg Asp Phe Pro Ile  
 50 55 60  
 Thr Trp Pro Trp Lys Trp Trp Lys Gly Gly Cys Glu Glu Leu Arg Asn  
 65 70 75 80  
 Glu Cys Cys Gln Leu Leu Gly Gln Met Pro Ser Glu Cys Arg Cys Asp  
 85 90 95  
 Ala Ile Trp Arg Ser Ile Gln Arg Glu Leu Gly Gly Phe Phe Gly Thr  
 100 105 110  
 Gln Gln Gly Leu Ile Gly Lys Arg Leu Lys Ile Ala Lys Ser Leu Pro  
 115 120 125  
 Thr Gln Ser Thr Trp Ala Leu Ser Ala Ile Ser Pro Asn Ser Met Val  
 130 135 140  
 Ser His Ile Ala Gly Lys Ser Ser Ile Leu Arg Ala Leu Pro Val Asp  
 145 150 155 160  
 Val Leu Ala Asn Ala Tyr Arg Ile Ser Arg Gln Glu Ala Arg Asn Leu  
 165 170 175  
 Lys Asn Asn Arg Gly Gln Glu Ser Gly Val Phe Thr Pro Lys Phe Thr  
 180 185 190  
 Gln Thr Ser Phe Gln Pro Tyr Pro Glu Gly Glu Asp Glu Ser Ser Leu  
 195 200 205  
 Ile Asn Lys Ala Ser Glu  
 210

&lt;210&gt; 109

&lt;211&gt; 1044

&lt;212&gt; DNA

&lt;213&gt; Hordeum vulgare

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1041)

&lt;223&gt; c-hordein

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (481)..(482)

&lt;223&gt; /transl\_except=(pos:481..483,aa:OTHER)

&lt;400&gt; 109

atg aag acg ttc ctc acc ttt gtc ctc ctt gcc atg gcg atg agc atc 48  
 Met Lys Thr Phe Leu Thr Phe Val Leu Leu Ala Met Ala Met Ser Ile  
 1 5 10 15

gtc act acc gct agg cag cta aac cct agc cac caa gag ttg caa tca 96  
 Val Thr Thr Ala Arg Gln Leu Asn Pro Ser His Gln Glu Leu Gln Ser  
 20 25 30

cca caa caa cca ttt ctg aaa caa caa tca tat ctg caa caa cca tat 144  
 Pro Gln Gln Pro Phe Leu Lys Gln Gln Ser Tyr Leu Gln Gln Pro Tyr  
 35 40 45

## 112

cca	caa	caa	cca	tat	cta	ccg	cag	caa	cca	ttc	ccc	aca	ccc	caa	caa	192
Pro	Gln	Gln	Pro	Tyr	Leu	Pro	Gln	Gln	Pro	Phe	Pro	Thr	Pro	Gln	Gln	
50						55					60					
ttt	ttc	ccc	tat	cta	cca	cag	caa	aca	ttt	ccc	cca	tcc	caa	caa	cca	240
Phe	Phe	Pro	Tyr	Leu	Pro	Gln	Gln	Thr	Phe	Pro	Pro	Ser	Gln	Gln	Pro	
65					70				75						80	
aac	ccc	cta	caa	cca	caa	caa	cca	ttc	ccc	ctg	caa	ccc	caa	cca	cca	288
Asn	Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	Pro	Pro	
			85					90						95		
caa	caa	cct	ttt	cct	cag	ccc	caa	caa	cca	aat	ccc	cag	caa	cca	caa	336
Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Gln	Pro	Asn	Pro	Gln	Gln	Pro	Gln	
			100					105					110			
caa	cct	ttc	ccc	cgg	caa	cca	caa	caa	ata	gta	ccc	cag	caa	cca	caa	384
Gln	Pro	Phe	Pro	Arg	Gln	Pro	Gln	Gln	Ile	Val	Pro	Gln	Gln	Pro	Gln	
			115					120					125			
caa	cca	ttc	cct	cag	caa	cca	caa	caa	cct	ttt	cct	cag	ccc	caa	caa	432
Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Gln	
			130					135					140			
cca	ttc	tct	tg	caa	cca	caa	caa	cca	ttt	ctc	cag	ccc	cta	caa	cta	480
Pro	Phe	Ser	Trp	Gln	Pro	Gln	Gln	Pro	Phe	Leu	Gln	Pro	Leu	Gln	Leu	
145						150				155					160	
tag	ccc	ctg	caa	gca	caa	caa	cca	ttc	ccc	ttg	caa	cct	caa	cta	cca	528
	Pro	Leu	Gln	Ala	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	Leu	Pro	
				165					170					175		
ttt	ccg	caa	ccc	caa	caa	cca	att	gga	cag	caa	cca	aaa	caa	cca	ctc	576
Phe	Pro	Gln	Pro	Gln	Gln	Pro	Ile	Gly	Gln	Gln	Pro	Lys	Gln	Pro	Leu	
			180						185					190		
ctg	cag	caa	cca	caa	caa	aca	att	ccc	cag	caa	cca	caa	caa	cca	ttc	624
Leu	Gln	Gln	Pro	Gln	Gln	Thr	Ile	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	
			195				200						205			
ccc	ctg	cag	ccg	caa	caa	cca	ttc	ccc	caa	caa	cca	caa	caa	cca	ctt	672
Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	
			210				215						220			
ccc	caa	caa	ccc	caa	caa	ata	att	tcc	cag	caa	ccc	caa	caa	cca	ttc	720
Pro	Gln	Gln	Pro	Gln	Gln	Ile	Ile	Ser	Gln	Gln	Pro	Gln	Gln	Pro	Phe	
225						230					235				240	
cct	cta	caa	cct	caa	caa	cca	ttc	ccc	caa	ccc	caa	cca	ttc	ccc	cag	768
Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Pro	Phe	Pro	Gln	
						245					250				255	
gag	caa	ccc	caa	caa	gca	ttc	ccc	cta	caa	ccg	caa	caa	cca	ttc	ccc	816
Glu	Gln	Pro	Gln	Gln	Ala	Phe	Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	
			260						265					270		
gag	gaa	tca	gaa	caa	ata	att	acc	caa	caa	cca	ttc	cct	cta	caa	cca	864
Glu	Glu	Ser	Glu	Gln	Ile	Ile	Thr	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	
			275				280						285			
caa	caa	ctg	ttc	ccc	cag	caa	cca	caa	caa	cca	ctt	ccc	cag	ccc	caa	912
Gln	Gln	Leu	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	Pro	Gln	Pro	Gln	
			290				295						300			
caa	cca	ttc	cgc	caa	cta	cca	aaa	tat	ata	att	ccc	cag	caa	cct	caa	960
Gln	Pro	Phe	Arg	Gln	Leu	Pro	Lys	Tyr	Ile	Ile	Pro	Gln	Gln	Pro	Gln	
305						310					315				320	

## 113

caa cca ttc ctt ctg caa cca cac caa cct cag caa cct tat gca caa 1008  
 Gln Pro Phe Leu Leu Gln Pro His Gln Pro Gln Gln Pro Tyr Ala Gln  
 325 330 335

caa gac atc tgg agt gat ata gcc ctc ttg ggc taa 1044  
 Gln Asp Ile Trp Ser Asp Ile Ala Leu Leu Gly  
 340 345

&lt;210&gt; 110

&lt;211&gt; 160

&lt;212&gt; PRT

&lt;213&gt; Hordeum vulgare

&lt;400&gt; 110

Met Lys Thr Phe Leu Thr Phe Val Leu Leu Ala Met Ala Met Ser Ile  
 1 5 10 15  
 Val Thr Thr Ala Arg Gln Leu Asn Pro Ser His Gln Glu Leu Gln Ser  
 20 25 30  
 Pro Gln Gln Pro Phe Leu Lys Gln Gln Ser Tyr Leu Gln Gln Pro Tyr  
 35 40 45  
 Pro Gln Gln Pro Tyr Leu Pro Gln Gln Pro Phe Pro Thr Pro Gln Gln  
 50 55 60  
 Phe Phe Pro Tyr Leu Pro Gln Gln Thr Phe Pro Pro Ser Gln Gln Pro  
 65 70 75 80  
 Asn Pro Leu Gln Pro Gln Gln Pro Phe Pro Leu Gln Pro Gln Pro Pro  
 85 90 95  
 Gln Gln Pro Phe Pro Gln Pro Gln Gln Pro Asn Pro Gln Gln Pro Gln  
 100 105 110  
 Gln Pro Phe Pro Arg Gln Pro Gln Gln Ile Val Pro Gln Gln Pro Gln  
 115 120 125  
 Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Phe Pro Gln Pro Gln Gln  
 130 135 140  
 Pro Phe Ser Trp Gln Pro Gln Gln Pro Phe Leu Gln Pro Leu Gln Leu  
 145 150 155 160

&lt;210&gt; 111

&lt;211&gt; 186

&lt;212&gt; PRT

&lt;213&gt; Hordeum vulgare

&lt;400&gt; 111

Pro Leu Gln Ala Gln Gln Pro Phe Pro Leu Gln Pro Gln Leu Pro Phe  
 1 5 10 15  
 Pro Gln Pro Gln Gln Pro Ile Gly Gln Gln Pro Lys Gln Pro Leu Leu  
 20 25 30  
 Gln Gln Pro Gln Gln Thr Ile Pro Gln Gln Pro Gln Gln Pro Phe Pro  
 35 40 45  
 Leu Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Leu Pro  
 50 55 60  
 Gln Gln Pro Gln Gln Ile Ile Ser Gln Gln Pro Gln Gln Pro Phe Pro  
 65 70 75 80  
 Leu Gln Pro Gln Gln Pro Phe Pro Gln Pro Gln Pro Phe Pro Gln Glu  
 85 90 95  
 Gln Pro Gln Gln Ala Phe Pro Leu Gln Pro Gln Gln Pro Phe Pro Glu  
 100 105 110  
 Glu Ser Glu Gln Ile Ile Thr Gln Gln Pro Phe Pro Leu Gln Pro Gln  
 115 120 125  
 Gln Leu Phe Pro Gln Gln Pro Gln Gln Pro Leu Pro Gln Pro Gln Gln  
 130 135 140  
 Pro Phe Arg Gln Leu Pro Lys Tyr Ile Ile Pro Gln Gln Pro Gln Gln  
 145 150 155 160



## 114

Pro Phe Leu Leu Gln Pro His Gln Pro Gln Gln Pro Tyr Ala Gln Gln  
 165 170 175  
 Asp Ile Trp Ser Asp Ile Ala Leu Leu Gly  
 180 185

&lt;210&gt; 112

&lt;211&gt; 924

&lt;212&gt; DNA

&lt;213&gt; Triticum aestivum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(921)

&lt;223&gt; glutenin-1D1

&lt;400&gt; 112

atg aag acc ttc ctc gtc ttt gcc ctc ctc gcc gtt gcg gcg aca agt	48
Met Lys Thr Phe Leu Val Phe Ala Leu Leu Ala Val Ala Ala Thr Ser	
1 5 10 15	
gca att gcg cag atg gag act aga tgc atc cct ggt ttg gag aga cca	96
Ala Ile Ala Gln Met Glu Thr Arg Cys Ile Pro Gly Leu Glu Arg Pro	
20 25 30	
tgg cag cag caa cca tta cca cca caa cag aca ttt cca caa caa cca	144
Trp Gln Gln Gln Pro Leu Pro Pro Gln Gln Thr Phe Pro Gln Gln Pro	
35 40 45	
cta ttt tca caa caa caa caa caa caa cta ttt cct caa caa cca tca	192
Leu Phe Ser Gln Gln Gln Gln Gln Gln Leu Phe Pro Gln Gln Pro Ser	
50 55 60	
ttt tcg cag caa caa cca cca ttt tgg cag caa caa cca cca ttt tct	240
Phe Ser Gln Gln Gln Pro Pro Phe Trp Gln Gln Gln Pro Pro Phe Ser	
65 70 75 80	
cag caa caa cca att cta cca cag caa cca cca ttt tcg cag caa caa	288
Gln Gln Gln Pro Ile Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln	
85 90 95	
caa cta gtt cta ccg caa caa cca cca ttt tca cag caa caa caa cca	336
Gln Leu Val Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln Gln Pro	
100 105 110	
gtt tta cct cca caa caa tca cct ttt cca caa caa caa caa caa cac	384
Val Leu Pro Pro Gln Gln Ser Pro Phe Pro Gln Gln Gln Gln Gln His	
115 120 125	
caa cag ctg gtg caa caa caa atc cct gtt gtt cag cca tcc att ttg	432
Gln Gln Leu Val Gln Gln Gln Ile Pro Val Val Gln Pro Ser Ile Leu	
130 135 140	
cag cag cta aac cca tgc aag gta ttc ctc cag cag cag tgc agc cct	480
Gln Gln Leu Asn Pro Cys Lys Val Phe Leu Gln Gln Gln Cys Ser Pro	
145 150 155 160	
gtg gca atg cca caa cgt ctt gct agg tcg caa atg ttg cag cag agc	528
Val Ala Met Pro Gln Arg Leu Ala Arg Ser Gln Met Leu Gln Gln Ser	
165 170 175	
agt tgc cat gtg atg caa caa caa tgt tgc cag cag ttg ccg caa atc	576
Ser Cys His Val Met Gln Gln Gln Cys Cys Gln Gln Leu Pro Gln Ile	
180 185 190	
ccc cag caa tcc cgc tat gag gca atc cgt gct atc atc tac tcc atc	624
Pro Gln Gln Ser Arg Tyr Glu Ala Ile Arg Ala Ile Ile Tyr Ser Ile	
195 200 205	

## 115

atc ctg caa gaa caa caa cag gtt cag ggt tcc atc caa tct cag cag 672  
 Ile Leu Gln Glu Gln Gln Val Gln Gly Ser Ile Gln Ser Gln Gln  
 210 215 220  
 cag caa ccc caa cag ttg ggc caa tgt gtt tcc caa ccc caa cag cag 720  
 Gln Gln Pro Gln Gln Leu Gly Gln Cys Val Ser Gln Pro Gln Gln Gln  
 225 230 235 240  
 tcg cag cag caa ctc ggg caa caa cct caa caa caa caa ttg gca cag 768  
 Ser Gln Gln Gln Leu Gly Gln Gln Pro Gln Gln Gln Gln Leu Ala Gln  
 245 250 255  
 ggt acc ttt ttg cag cca cac cag ata gct cag ctt gag gtg atg act 816  
 Gly Thr Phe Leu Gln Pro His Gln Ile Ala Gln Leu Glu Val Met Thr  
 260 265 270  
 tcc att gcg ctc cgt atc ctg cca acg atg tgc agt gtt aat gtg ccg 864  
 Ser Ile Ala Leu Arg Ile Leu Pro Thr Met Cys Ser Val Asn Val Pro  
 275 280 285  
 ttg tac aga acc acc act agt gtg cca ttc ggc gtt ggc acc gga gtt 912  
 Leu Tyr Arg Thr Thr Thr Ser Val Pro Phe Gly Val Gly Thr Gly Val  
 290 295 300  
 ggt gcc tac tga 924  
 Gly Ala Tyr  
 305

&lt;210&gt; 113

&lt;211&gt; 307

&lt;212&gt; PRT

&lt;213&gt; Triticum aestivum

&lt;400&gt; 113

Met Lys Thr Phe Leu Val Phe Ala Leu Leu Ala Val Ala Ala Thr Ser  
 1 5 10 15  
 Ala Ile Ala Gln Met Glu Thr Arg Cys Ile Pro Gly Leu Glu Arg Pro  
 20 25 30  
 Trp Gln Gln Gln Pro Leu Pro Pro Gln Gln Thr Phe Pro Gln Gln Pro  
 35 40 45  
 Leu Phe Ser Gln Gln Gln Gln Gln Gln Leu Phe Pro Gln Gln Pro Ser  
 50 55 60  
 Phe Ser Gln Gln Gln Pro Pro Phe Trp Gln Gln Gln Pro Pro Phe Ser  
 65 70 75 80  
 Gln Gln Gln Pro Ile Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln  
 85 90 95  
 Gln Leu Val Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln Gln Pro  
 100 105 110  
 Val Leu Pro Pro Gln Gln Ser Pro Phe Pro Gln Gln Gln Gln His  
 115 120 125  
 Gln Gln Leu Val Gln Gln Gln Ile Pro Val Val Gln Pro Ser Ile Leu  
 130 135 140  
 Gln Gln Leu Asn Pro Cys Lys Val Phe Leu Gln Gln Gln Cys Ser Pro  
 145 150 155 160  
 Val Ala Met Pro Gln Arg Leu Ala Arg Ser Gln Met Leu Gln Gln Ser  
 165 170 175  
 Ser Cys His Val Met Gln Gln Gln Cys Cys Gln Gln Leu Pro Gln Ile  
 180 185 190

## 116

Pro Gln Gln Ser Arg Tyr Glu Ala Ile Arg Ala Ile Ile Tyr Ser Ile  
 195 200 205

Ile Leu Gln Glu Gln Gln Gln Val Gln Gly Ser Ile Gln Ser Gln Gln  
 210 215 220

Gln Gln Pro Gln Gln Leu Gly Gln Cys Val Ser Gln Pro Gln Gln Gln  
 225 230 235 240

Ser Gln Gln Gln Leu Gly Gln Gln Pro Gln Gln Gln Gln Leu Ala Gln  
 245 250 255

Gly Thr Phe Leu Gln Pro His Gln Ile Ala Gln Leu Glu Val Met Thr  
 260 265 270

Ser Ile Ala Leu Arg Ile Leu Pro Thr Met Cys Ser Val Asn Val Pro  
 275 280 285

Leu Tyr Arg Thr Thr Thr Ser Val Pro Phe Gly Val Gly Thr Gly Val  
 290 295 300

Gly Ala Tyr  
 305

&lt;210&gt; 114

&lt;211&gt; 8482

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: binary  
expression vector

&lt;400&gt; 114

```

ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttattctaataa aaacgctctt ttctcttagg tttaccgcc aatatacct gtcaaact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aatcaccact ttgtacaaga aagctgggtc tagatgacgg acaatcagta 240
aattgaacgg agaattattat tcataaaaaat acgatagtaa cgggtgatattt cattataga 300
atgaaccgaa accggcggta aggatctgag ctacacatgc tcagggtttt tacaacgtgc 360
acaacagaat tgaaagcaaa tatcatgcga tcataggcgt ctgcataatc tcattaaagc 420
aggaggcctt ctagactgca ggcgccgcc caccgcggtg ggctggctat gaagaaatta 480
taatcgtgta aaacttagtg agtgtgtatg aatgaaagta ttgcaaaatc ctcattatat 540
agactacatg cataactagt tgcattgtaa tttgtagttt tcttcattat tgcattctcc 600
aagtggatgt catggtttta cacatggcct ccatgcaaat catttcctaa atatttttaa 660
actttccaca gggcatccat gcatgcacct caaaacttgt gtgtggtaac attgttgtct 720
tgaaaaatta ctaaactttt tgtccacgtg acgttcacgc acctcaaact ttgtgtggta 780
ccattattat cctcaagaat tattgaatgt ttggtgtata tgccatccat gcagcattgc 840
aacaattaaa tctccaaacc ttgtggtacc atattcactc actttaattc tcctatagta 900
gaaatattag caaatattta catttccagt tgattagtat atgtatttag aagacaaaaa 960
taatttagaa tcaattaatc aacttgcaaa ttgctaagtg ttggcaaacg ttagcataaa 1020
agggtgtata aatttagtac caaatataaa aatttatcgc aaatcaaata cataacacac 1080
atagtaaaac aaaaacaaat tacaagggtt tagacgttta gtggcaatgt gtaaatttgc 1140
tcgactgaat tggttccttt aagcctgctt tttgtacaa acttgtgata attcactggc 1200
cgctgtttta caacgactca ggatcctgtc aaacactgat agtttaaact gaaggcggga 1260
aacgacaatc tgatcatgag cggagaatta agggagtcac gttatgacct ccgccgatga 1320
cgcggggaca gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc 1380
agccgcgggt ttctggagtt taatgagcta agcacatacg tcagaaacca ttattgcgcg 1440
ttcaaaagtc gcctaaggtc actatcagct agcaaataat tcttgtcaaa aatgctccac 1500
tgacgttcca taaattcccc tcggtatcca attagagtct catattcact ctcaatccaa 1560
ataatctgca ccggtatctg atcgtttcgc atgattgaac aagatggatt gcacgcagg 1620
tctccggccg cttgggtgga gaggtatttc ggctatgact gggcacaaca gacaatcggc 1680

```

tgctctgatg	ccgccgtggt	ccggctgtca	gcgcaggggc	gcccggttct	ttttgtcaag	1740
accgacctgt	ccggtgccct	gaatgaactg	caggacgagg	cagcgcggct	atcgtggctg	1800
gccacgacgg	gcgttccttg	cgcagctgtg	ctcgacgttg	tcactgaagc	gggaaggagc	1860
tggctgctat	tgggcgaagt	gccggggcag	gatctcctgt	catctcacct	tgctcctgcc	1920
gagaaagtat	ccatcatggc	tgatgcaatg	cggcggctgc	atacgttga	tccggctacc	1980
tgcccatcgc	accaccaagc	gaaacatcgc	atcgagcgag	cacgtactcg	gatggaagcc	2040
ggtcttgcgc	atcaggatga	tctggacgaa	gagcatcagg	ggctcgcgcc	agccgaactg	2100
ttcgccaggc	tcaaggcgcg	catgcccgcg	ggcgaggatc	tcgtcgtgac	ccatggcgat	2160
gcctgcttgc	cgaatatcat	ggtggaaaat	ggcgcgtttt	ctggattcat	cgactgtggc	2220
cggctgggtg	tggcggaccg	ctatcaggac	atagcgttgg	ctaccctgta	tattgctgaa	2280
gagcttggcg	gcgaatgggc	tgaccgcttc	ctcgtgcttt	acggtatcgc	cgctcccgat	2340
tcgcagcgca	tcgccttcta	tcgccttctt	gacgagttct	tctgagcggg	acceaaagctc	2400
tagatcttgc	tgcgttcgga	tattttcgtg	gagttcccgc	cacagaccgc	gatgatcccc	2460
gatcgttcaa	acatttggca	ataaagtttc	ttaagattga	atcctgttgc	cggctcttgcg	2520
atgattatca	tataatttct	gttgaattac	gttaagcatg	taataattaa	catgtaatgc	2580
atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	cgcaattata	catttaatac	2640
gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	tatcgcgcg	ggtgtcatct	2700
atgttactag	atcgggcctc	ctgtcaagct	ctgcttggta	ataattgtca	ttagattgtt	2760
tttatgcata	gatgcactcg	aaatcagcca	atttttagaca	agtatcaaac	ggatgttaat	2820
tcagtacatt	aaagacgtcc	gcaatgtgtt	attaagttgt	ctaagcgtca	atttgtttac	2880
accacaatat	atcctgccac	cagccagcca	acagctcccc	gaccggcagc	tcggcacaaa	2940
atcaccacgc	gttaccacca	cgcgcggcgg	cgcgatggtg	ttgaccgtgt	tcgcggcat	3000
tgccgagttc	gagcgttccc	taatcatcga	cgcacccgg	agcgggcgcg	agggcgccaa	3060
ggccccaggc	gtgaagtttg	gccccgcgcc	tacctcacc	ccggcacaga	tcgcgcacgc	3120
ccgcgagctg	atcgaccagg	aaggccgcac	cgtgaaagag	gcggctgcac	tgcttggcgt	3180
gcacgcctcg	accctgtacc	gcgcacttga	gcgcagcgag	gaagtgacgc	ccaccgaggc	3240
caggcggcgc	ggtgccttcc	gtgaggacgc	attgaccgag	gccgacgccc	tgggcgccgc	3300
cgagaatgaa	gcgaagagg	aacaagcatg	aaaccgcacc	aggacggcca	ggacgaaccg	3360
tttttcatta	ccgaagagat	cgaggcggag	atgatcgcg	ccgggtacgt	gttcgagccg	3420
ccgcgcacgc	tctcaaccgt	gcggctgcat	gaaatcctgg	ccggtttgtc	tgatgccaa	3480
ctggcgccct	ggccggccag	cttggccgct	gaagaaaccg	agcgcgcgcg	tctaaaaagg	3540
tgatgtgtat	ttgagtaaaa	cagcttgctg	catgcggtcg	ctgcgtatat	gatgcgatga	3600
gtaataaac	aaatacgcaa	ggggaacgca	tgaaggttat	cgctgtactt	aaccagaaa	3660
gcgggtcagg	caagacgacc	atcgcaaccc	atctagcccc	cgccctgcaa	ctcgccgggg	3720
ccgatgttct	gttagtcgat	tccgatcccc	agggcagtg	ccgcgattgg	gcggccgtgc	3780
gggaagatca	accgctaacc	gttgtcggca	tcgaccgccc	gacgattgac	cgcgacgtga	3840
aggccatcgg	ccggcgcgac	ttcgtagtga	tcgacggagc	gccccaggcg	gcggacttgg	3900
ctgtgtccgc	gatcaaggca	gccgacttgc	tgctgattcc	ggtgcagcca	agcccttacg	3960
acatatgggc	caccgcccgc	ctggtggagc	tggttaagca	gcgcattgag	gtcacggatg	4020
gaaggctaca	agcggccttt	gtcgtgtcgc	gggcgatcaa	aggcacgcgc	atcggcggtg	4080
aggttgcga	ggcgctggcc	gggtacgagc	tgccattct	tgagtcccg	atcacgcagc	4140
gcgtgagcta	cccaggcact	gccgcgcgcg	gcacaaccgt	tcttgaatca	gaaccgagg	4200
gcgacgctgc	ccgcgaggtc	caggcgctgg	ccgctgaaat	taaataaaaa	ctcattttgag	4260
ttaatgaggt	aaagagaaaa	tgagcaaaa	cacaaacacg	ctaagtgccg	gccgtccgag	4320
cgcacgcagc	agcaaggctg	caacgttggc	cagcctggca	gacacgccag	ccatgaagcg	4380
ggtcaacttt	cagttgccgg	cggaggatca	caccaagctg	aagatgtacg	cggtacgcca	4440
aggcaagacc	attaccgagc	tgctatctga	atacatcgcg	cagctaccag	agtaaatgag	4500
caaatgaata	aatgagtaga	tgaattttag	cggctaaagg	aggcggcatg	gaaaatcaag	4560
aacaaccagg	caccgacgcc	gtggaatgcc	ccatgtgtgg	aggaaacggc	ggttggccag	4620
gcgtaagcgg	ctgggttgtc	tgccggccct	gcaatggcac	tggaaccccc	aagcccagg	4680
aatcggcgtg	agcggctcga	aaccatccgg	cccggtacaa	atcggcgcg	cgctgggtga	4740
tgacctgggtg	gagaagttga	aggccgcgca	ggccgcccag	cggcaacgca	tcgaggcgca	4800
agcacgcccc	ggtgaatcgt	ggcaagcggc	cgctgatcga	atccgcaaag	aatcccggca	4860
accgcgggca	gccggtgcgc	cgtcgattag	gaagccgccc	aagggcgacg	agcaaccaga	4920
ttttttcgtt	ccgatgctct	atgacgtggg	caccgcgat	agtcgcagca	tcatggacgt	4980
ggccgttttc	cgtctgtcga	agcgtgaccg	acgagctggc	gaggtgatcc	gctacgagct	5040
tccagacggg	cacgtagagg	tttccgcagg	gccggccggc	atggccagtg	tgtgggatta	5100

cgacctggta ctgatggcgg tttcccatct aaccgaatcc atgaaccgat accgggaagg 5160  
 gaagggagac aagcccgccc gcgtgttccg tccacacgtt gcggacgtac tcaagttctg 5220  
 ccggcgagcc gatggcggaa agcagaaaga cgacctggta gaaacctgca ttcggtaaa 5280  
 caccacgcac gttgccatgc agcgtacgaa gaaggccaag aacggccgcc tggtagcgg 5340  
 atccgagggg gaagccttga ttagccgcta caagatcgta aagagcgaaa ccgggcggcc 5400  
 ggagtacatc gagatcgagc tagctgattg gatgtaccgc gagatcacag aaggcaagaa 5460  
 cccggacgtg ctgacggttc accccgatta ctttttgatc gatcccgga tcggccgttt 5520  
 tctctaccgc ctggcacgcc gcgccgcagg caaggcagaa gccagatggg tgttcaagac 5580  
 gatctacgaa cgcagtggca gcgccggaga gttcaagaag ttctgtttca ccgtgcgcaa 5640  
 gctgatccgg tcaaatgacc tgccggagta cgatttgaag gaggaggcgg ggcaggctgg 5700  
 cccgatcccta gtcatgcgt accgcaacct gatcgagggc gaagcatccg ccggttcccta 5760  
 atgtacggag cagatgctag ggcaaatg cctagcaggg gaaaaaggtc gaaaaggctc 5820  
 ctttccctgt gatagcacgt acattgggaa cccaaagccg tacattggga accggaaccc 5880  
 gtacattggg aacccaaagc cgtacattgg gaaccggtca cacatgtaag tgactgatat 5940  
 aaaagagaaa aaaggcgatt tttccgcta aaactcttta aaacttatta aaactcttaa 6000  
 aaccgcctg gcctgtgcat aactgtctgg ccagcgcaca gccgaagagc tgcaaaaagc 6060  
 gcctaccctt cggctgcgtg gctccctacg ccccgccgct tcgctcgcc ctatcgccgc 6120  
 cgctggccgc tcaaaaatgg ctggcctacg gccaggcaat ctaccagggc gcggacaagc 6180  
 cgcgcgctc ccactcgacc gccggcgccc acatcaaggc accctgcctc gcgcttctg 6240  
 gtgatgacgg tgaaaacctc tgacacatgc agctcccga gacggtcaca gcttgtctgt 6300  
 aagcggatgc cgggagcaga caagcccgtc agggcgcgct agcgggtgtt ggccgggtgtc 6360  
 gggcgcgac catgacccag tcacgtagcg atagcggagt gtatactggc ttaactatgc 6420  
 ggcatcagag cagattgtac tgagagtga ccatatgcgg tgtgaaatac cgcacagatg 6480  
 cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg 6540  
 ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc 6600  
 cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 6660  
 gaaccgtaaa aaggccgctg tgctggcgtt tttccatagg ctccgcccc ctgacgagca 6720  
 tcacaaaaat cgacgctcaa gtcagagggt gcgaaacccg acaggactat aaagatacca 6780  
 gggttttccc cctggaagct cctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 6840  
 atacctgtcc gcctttctcc cttcggaag cgtggcgctt tctcatagct cactgtgtag 6900  
 gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt 6960  
 tcagcccgac cgctgcgct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 7020  
 cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 7080  
 cgggtgtaca gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt 7140  
 tggatctgc gctctgtga agccagttac tttcggaaaa agagtgtgta gctcttgatc 7200  
 cggcaaaaca accaccgctg gtagcgggtg ttttttggg tgcaagcagc agattacgcg 7260  
 cagaaaaaaa ggatctcaag aagatccttt gatctttct acgggtctg acgctcagt 7320  
 gaacgaaaac tcacgttaag ggattttgg catgcatgat atatctccca atttgtgtag 7380  
 ggcttattat gcacgttaa aaataataaa agcagacttg acctgatagt ttggctgtga 7440  
 gcaattatgt gcttagtgca tctaacgctt gagttaagcc gcgccgcgaa gcggcgctcg 7500  
 cttgaacgaa tttctagcta gacattatt gccgactacc ttgggtgatc cgctttcac 7560  
 gtagtgga aattcttcca actgatctgc gcgcgaggcc aagcgatctt cttcttgtcc 7620  
 aagataagcc tgtctagct caagtatgac gggctgatac tgggcccggc ggcgtccat 7680  
 tgcccagtcg gcagcgacat ccttcggcgc gattttgccc gttactgcgc tgtaccaa 7740  
 gcgggacaa gtaagcacta catttcgctc atcgccagcc cagtcgggcg gcgagttcca 7800  
 tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg gatcaaagag 7860  
 ttctccgccc gctggaccta ccaaggcaac gctatgttct cttgcttttg tcagcaagat 7920  
 agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt cattgcgtg 7980  
 ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga tgtcgtcgtg 8040  
 cacaacaatg gtgacttcta cagcgcgag aatctcgctc tctccagggg aagccgaagt 8100  
 ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta cggtcaccgt 8160  
 aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg agccgtacaa 8220  
 atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta cctctgatag 8280  
 ttgagtcgat acttcggcga tcaccgcttc ccccatgatg tttaactttg ttttagggcg 8340  
 actgccctgc tgcgtaacat ogttgctgct ccataacatc aaacatcgac ccacggcgta 8400  
 acgcgcttgc tgcttggatg cccgaggcat agactgtacc ccaaaaaaac agtcataaca 8460  
 agccatgaaa accgccactg cg 8482

```

<210> 115
<211> 575
<212> DNA
<213> Brassica napus
<220>
<221> misc_feature
<222> (1)..(6)
<223> restriction site

<220>
<221> misc_feature
<222> (570)..(575)
<223> restriction site

<220>
<221> misc_feature
<222> (7)..(575)
<223> coding for homogentisate-1,2-dioxygenase (HDG)

<400> 115
gtcgacgggc cgatgggggc gaaggggtctt gctgcaccaa gagattttct tgcaccaacg 60
gcatggtttg aggaagggct acggcctgac tacactattg ttcagaagtt tggcgggtgaa 120
ctctttactg ctaaacaaga tttctctccg ttcaatgtgg ttgcctggca tggcaattac 180
gtgccttata agtatgacct gcacaagttc tgtccataca aactgtcctt tgtagaccat 240
ggagatccat ctgtaaatac agttctgaca gcaccaacgg ataaacctgg tgtggccttg 300
cttgattttg tcatattccc tctcgtttgg ttggttgctg agcatacctt tcgacctcct 360
tactaccatc gtaactgcat gagtgaattt atgggcctaa tctatggtgc ttacgaggcc 420
aaagctgatg gatttctacc tgggtggcgca agtcttcaca gttgtatgac acctcatggt 480
ccagatacaa ccacatacga ggcgacgatt gctcgtgtaa atgcaatggc tccttataag 540
ctcacaggca ccatggcctt catgtttgag gtacc 575

<210> 116
<211> 1386
<212> DNA
<213> Arabidopsis thaliana
<220>
<221> CDS
<222> (1)..(1383)
<223> coding for homogentisate-1,2-dioxygenase (HDG)

<400> 116
atg gaa gag aag aag aag gag ctt gaa gag ttg aag tat caa tca ggt 48
Met Glu Glu Lys Lys Lys Glu Leu Glu Glu Leu Lys Tyr Gln Ser Gly
1 5 10 15
ttt ggt aac cac ttc tca tcg gaa gca atc gcc gga gct tta ccg tta 96
Phe Gly Asn His Phe Ser Ser Glu Ala Ile Ala Gly Ala Leu Pro Leu
20 25 30
gat cag aac agt cct ctt ctt tgt cct tac ggt ctt tac gcc gaa cag 144
Asp Gln Asn Ser Pro Leu Leu Cys Pro Tyr Gly Leu Tyr Ala Glu Gln
35 40 45
atc tcc ggt act tct ttc act tct cct cgc aag ctc aat caa aga agt 192
Ile Ser Gly Thr Ser Phe Thr Ser Pro Arg Lys Leu Asn Gln Arg Ser
50 55 60
tgg ttg tac cgg gtt aaa cca tcg gtt aca cat gaa ccg ttc aag cct 240
Trp Leu Tyr Arg Val Lys Pro Ser Val Thr His Glu Pro Phe Lys Pro
65 70 75 80
cgt gta cca gct cat aag aag ctt gtg agt gag ttt gat gca tca aat 288
Arg Val Pro Ala His Lys Lys Leu Val Ser Glu Phe Asp Ala Ser Asn
85 90 95

```

## 120

agt cgt acg aat ccg act cag ctt cgg tgg aga cct gag gat att cct	336
Ser Arg Thr Asn Pro Thr Gln Leu Arg Trp Arg Pro Glu Asp Ile Pro	
100 105 110	
gat tcg gag att gat ttc gtt gat ggg tta ttt acc att tgt gga gct	384
Asp Ser Glu Ile Asp Phe Val Asp Gly Leu Phe Thr Ile Cys Gly Ala	
115 120 125	
gga agc tcg ttt ctt cgc cat ggc ttc gct att cac atg tat gtg gct	432
Gly Ser Ser Phe Leu Arg His Gly Phe Ala Ile His Met Tyr Val Ala	
130 135 140	
aac aca gga atg aaa gac tcc gca ttt tgc aac gct gat ggt gac ttc	480
Asn Thr Gly Met Lys Asp Ser Ala Phe Cys Asn Ala Asp Gly Asp Phe	
145 150 155 160	
ttg tta gtt cct caa aca gga agg cta tgg att gaa act gag tgt gga	528
Leu Leu Val Pro Gln Thr Gly Arg Leu Trp Ile Glu Thr Glu Cys Gly	
165 170 175	
agg ctt ttg gta act cct ggt gag att gct gtt ata cca caa ggt ttc	576
Arg Leu Leu Val Thr Pro Gly Glu Ile Ala Val Ile Pro Gln Gly Phe	
180 185 190	
cgt ttc tcc ata gat tta ccg gat ggg aag tct cgt ggt tat gtt gct	624
Arg Phe Ser Ile Asp Leu Pro Asp Gly Lys Ser Arg Gly Tyr Val Ala	
195 200 205	
gaa atc tat ggg gct cat ttt cag ctt cct gat ctt gga cca ata ggt	672
Glu Ile Tyr Gly Ala His Phe Gln Leu Pro Asp Leu Gly Pro Ile Gly	
210 215 220	
gct aat ggt ctt gct gca tca aga gat ttt ctt gca cca aca gca tgg	720
Ala Asn Gly Leu Ala Ala Ser Arg Asp Phe Leu Ala Pro Thr Ala Trp	
225 230 235 240	
ttt gag gat gga ttg cgg cct gaa tac aca att gtt cag aag ttt ggc	768
Phe Glu Asp Gly Leu Arg Pro Glu Tyr Thr Ile Val Gln Lys Phe Gly	
245 250 255	
ggt gaa ctc ttt act gct aaa caa gat ttc tct cca ttc aat gtg gtt	816
Gly Glu Leu Phe Thr Ala Lys Gln Asp Phe Ser Pro Phe Asn Val Val	
260 265 270	
gcc tgg cat ggc aat tac gtg cct tat aag tat gac ctg aag aag ttc	864
Ala Trp His Gly Asn Tyr Val Pro Tyr Lys Tyr Asp Leu Lys Lys Phe	
275 280 285	
tgt cca tac aac act gtg ctt tta gat cat gga gat cca tct ata aat	912
Cys Pro Tyr Asn Thr Val Leu Leu Asp His Gly Asp Pro Ser Ile Asn	
290 295 300	
aca gtc ctt aca gca cca act gat aaa cct ggt gtg gcc ttg ctt gat	960
Thr Val Leu Thr Ala Pro Thr Asp Lys Pro Gly Val Ala Leu Leu Asp	
305 310 315 320	
ttt gtc ata ttt cct cct cga tgg ttg gtt gct gag cat act ttt cga	1008
Phe Val Ile Phe Pro Pro Arg Trp Leu Val Ala Glu His Thr Phe Arg	
325 330 335	
cct cct tac tat cat cgt aac tgc atg agt gaa ttt atg ggc tta atc	1056
Pro Pro Tyr Tyr His Arg Asn Cys Met Ser Glu Phe Met Gly Leu Ile	
340 345 350	
tac ggt gca tac gag gcg aaa gct gat gga ttt ctc cct ggc ggt gca	1104
Tyr Gly Ala Tyr Glu Ala Lys Ala Asp Gly Phe Leu Pro Gly Gly Ala	
355 360 365	

## 121

agt ctt cat agc tgt atg aca cct cat ggt cca gat act acc acg tac 1152  
 Ser Leu His Ser Cys Met Thr Pro His Gly Pro Asp Thr Thr Thr Tyr  
 370 375 380  
 gag gcg aca att gct cga gta aat gca atg gct cct tct aaa ctc aca 1200  
 Glu Ala Thr Ile Ala Arg Val Asn Ala Met Ala Pro Ser Lys Leu Thr  
 385 390 395 400  
 ggt acg atg gct ttc atg ttc gaa tca gca ttg atc cct aga gtc tgt 1248  
 Gly Thr Met Ala Phe Met Phe Glu Ser Ala Leu Ile Pro Arg Val Cys  
 405 410 415  
 cat tgg gct ctg gag tct cct ttc ctg gat cac gac tac tac cag tgt 1296  
 His Trp Ala Leu Glu Ser Pro Phe Leu Asp His Asp Tyr Tyr Gln Cys  
 420 425 430  
 tgg att ggc ctc aag tct cat ttc tcg cgc ata agc ttg gac aag aca 1344  
 Trp Ile Gly Leu Lys Ser His Phe Ser Arg Ile Ser Leu Asp Lys Thr  
 435 440 445  
 aat gtt gaa tca aca gag aaa gaa cca gga gct tcg gag taa 1386  
 Asn Val Glu Ser Thr Glu Lys Glu Pro Gly Ala Ser Glu  
 450 455 460

&lt;210&gt; 117

&lt;211&gt; 461

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 117

Met Glu Glu Lys Lys Lys Glu Leu Glu Glu Leu Lys Tyr Gln Ser Gly  
 1 5 10 15  
 Phe Gly Asn His Phe Ser Ser Glu Ala Ile Ala Gly Ala Leu Pro Leu  
 20 25 30  
 Asp Gln Asn Ser Pro Leu Leu Cys Pro Tyr Gly Leu Tyr Ala Glu Gln  
 35 40 45  
 Ile Ser Gly Thr Ser Phe Thr Ser Pro Arg Lys Leu Asn Gln Arg Ser  
 50 55 60  
 Trp Leu Tyr Arg Val Lys Pro Ser Val Thr His Glu Pro Phe Lys Pro  
 65 70 75 80  
 Arg Val Pro Ala His Lys Lys Leu Val Ser Glu Phe Asp Ala Ser Asn  
 85 90 95  
 Ser Arg Thr Asn Pro Thr Gln Leu Arg Trp Arg Pro Glu Asp Ile Pro  
 100 105 110  
 Asp Ser Glu Ile Asp Phe Val Asp Gly Leu Phe Thr Ile Cys Gly Ala  
 115 120 125  
 Gly Ser Ser Phe Leu Arg His Gly Phe Ala Ile His Met Tyr Val Ala  
 130 135 140  
 Asn Thr Gly Met Lys Asp Ser Ala Phe Cys Asn Ala Asp Gly Asp Phe  
 145 150 155 160  
 Leu Leu Val Pro Gln Thr Gly Arg Leu Trp Ile Glu Thr Glu Cys Gly  
 165 170 175  
 Arg Leu Leu Val Thr Pro Gly Glu Ile Ala Val Ile Pro Gln Gly Phe  
 180 185 190  
 Arg Phe Ser Ile Asp Leu Pro Asp Gly Lys Ser Arg Gly Tyr Val Ala  
 195 200 205



## 122

Glu Ile Tyr Gly Ala His Phe Gln Leu Pro Asp Leu Gly Pro Ile Gly  
 210 215 220  
 Ala Asn Gly Leu Ala Ala Ser Arg Asp Phe Leu Ala Pro Thr Ala Trp  
 225 230 235 240  
 Phe Glu Asp Gly Leu Arg Pro Glu Tyr Thr Ile Val Gln Lys Phe Gly  
 245 250 255  
 Gly Glu Leu Phe Thr Ala Lys Gln Asp Phe Ser Pro Phe Asn Val Val  
 260 265 270  
 Ala Trp His Gly Asn Tyr Val Pro Tyr Lys Tyr Asp Leu Lys Lys Phe  
 275 280 285  
 Cys Pro Tyr Asn Thr Val Leu Leu Asp His Gly Asp Pro Ser Ile Asn  
 290 295 300  
 Thr Val Leu Thr Ala Pro Thr Asp Lys Pro Gly Val Ala Leu Leu Asp  
 305 310 315 320  
 Phe Val Ile Phe Pro Pro Arg Trp Leu Val Ala Glu His Thr Phe Arg  
 325 330 335  
 Pro Pro Tyr Tyr His Arg Asn Cys Met Ser Glu Phe Met Gly Leu Ile  
 340 345 350  
 Tyr Gly Ala Tyr Glu Ala Lys Ala Asp Gly Phe Leu Pro Gly Gly Ala  
 355 360 365  
 Ser Leu His Ser Cys Met Thr Pro His Gly Pro Asp Thr Thr Thr Tyr  
 370 375 380  
 Glu Ala Thr Ile Ala Arg Val Asn Ala Met Ala Pro Ser Lys Leu Thr  
 385 390 395 400  
 Gly Thr Met Ala Phe Met Phe Glu Ser Ala Leu Ile Pro Arg Val Cys  
 405 410 415  
 His Trp Ala Leu Glu Ser Pro Phe Leu Asp His Asp Tyr Tyr Gln Cys  
 420 425 430  
 Trp Ile Gly Leu Lys Ser His Phe Ser Arg Ile Ser Leu Asp Lys Thr  
 435 440 445  
 Asn Val Glu Ser Thr Glu Lys Glu Pro Gly Ala Ser Glu  
 450 455 460

&lt;210&gt; 118

&lt;211&gt; 815

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (37)..(705)

&lt;223&gt; coding for maleylacetoacetate isomerase (MAAI)

&lt;400&gt; 118

gtaatctccg aagaagaaca aattccttgc tgaatc atg tct tat gtt acc.gat 54

Met Ser Tyr Val Thr Asp

1

5

ttt tat cag gcg aag ttg aag ctc tac tct tac tgg aga agc tca tgt 102

Phe Tyr Gln Ala Lys Leu Lys Leu Tyr Ser Tyr Trp Arg Ser Ser Cys

10

15

20

## 123

gct cat cgc gtc cgt atc gcc ctc act tta aaa ggg ctt gat tat gaa 150  
 Ala His Arg Val Arg Ile Ala Leu Thr Leu Lys Gly Leu Asp Tyr Glu  
 25 30 35  
 tat ata ccg gtt aat ttg ctc aaa ggg gat caa tcc gat tca gat ttc 198  
 Tyr Ile Pro Val Asn Leu Leu Lys Gly Asp Gln Ser Asp Ser Asp Phe  
 40 45 50  
 aag aag atc aat cca atg ggc act gta cca gcg ctt gtt gat ggt gat 246  
 Lys Lys Ile Asn Pro Met Gly Thr Val Pro Ala Leu Val Asp Gly Asp  
 55 60 65 70  
 gtt gtg att aat gac tct ttc gca ata ata atg tac ctg gat gat aag 294  
 Val Val Ile Asn Asp Ser Phe Ala Ile Ile Met Tyr Leu Asp Asp Lys  
 75 80 85  
 tat ccg gag cca ccg ctg tta cca agt gac tac cat aaa cgg gcg gta 342  
 Tyr Pro Glu Pro Pro Leu Leu Pro Ser Asp Tyr His Lys Arg Ala Val  
 90 95 100  
 aat tac cag gcg acg agt att gtc atg tct ggt ata cag cct cat caa 390  
 Asn Tyr Gln Ala Thr Ser Ile Val Met Ser Gly Ile Gln Pro His Gln  
 105 110 115  
 aat atg gct ctt ttt agg tat ctc gag gac aag ata aat gct gag gag 438  
 Asn Met Ala Leu Phe Arg Tyr Leu Glu Asp Lys Ile Asn Ala Glu Glu  
 120 125 130  
 aaa act gct tgg att act aat gct atc aca aaa gga ttc aca gct ctc 486  
 Lys Thr Ala Trp Ile Thr Asn Ala Ile Thr Lys Gly Phe Thr Ala Leu  
 135 140 145 150  
 gag aaa ctg ttg gtg agt tgc gct gga aaa tac gcg act ggt gat gaa 534  
 Glu Lys Leu Leu Val Ser Cys Ala Gly Lys Tyr Ala Thr Gly Asp Glu  
 155 160 165  
 gtt tac ttg gct gat ctt ttc cta gca cca cag atc cac gca gca ttc 582  
 Val Tyr Leu Ala Asp Leu Phe Leu Ala Pro Gln Ile His Ala Ala Phe  
 170 175 180  
 aac aga ttc cat att aac atg gaa cca ttc ccg act ctt gca agg ttt 630  
 Asn Arg Phe His Ile Asn Met Glu Pro Phe Pro Thr Leu Ala Arg Phe  
 185 190 195  
 tac gag tca tac aac gaa ctg cct gca ttt caa aat gca gtc ccg gag 678  
 Tyr Glu Ser Tyr Asn Glu Leu Pro Ala Phe Gln Asn Ala Val Pro Glu  
 200 205 210  
 aag caa cca gat act cct tcc acc atc tgattctgtg aaccgtaagc 725  
 Lys Gln Pro Asp Thr Pro Ser Thr Ile  
 215 220  
 ttctctcagt ctcagctcaa taaaatctct taggaaacaa caacaacacc ttgaacttaa 785  
 atgtatcata tgaaccagtt tacaaataat 815  
 <210> 119  
 <211> 223  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 119  
 Met Ser Tyr Val Thr Asp Phe Tyr Gln Ala Lys Leu Lys Leu Tyr Ser  
 1 5 10 15  
 Tyr Trp Arg Ser Ser Cys Ala His Arg Val Arg Ile Ala Leu Thr Leu  
 20 25 30

## 124

Lys Gly Leu Asp Tyr Glu Tyr Ile Pro Val Asn Leu Leu Lys Gly Asp  
                   35                                  40                                  45  
 Gln Ser Asp Ser Asp Phe Lys Lys Ile Asn Pro Met Gly Thr Val Pro  
                   50                                  55                                  60  
 Ala Leu Val Asp Gly Asp Val Val Ile Asn Asp Ser Phe Ala Ile Ile  
                   65                                  70                                  75                                  80  
 Met Tyr Leu Asp Asp Lys Tyr Pro Glu Pro Pro Leu Leu Pro Ser Asp  
                                   85                                  90                                  95  
 Tyr His Lys Arg Ala Val Asn Tyr Gln Ala Thr Ser Ile Val Met Ser  
                                   100                                  105                                  110  
 Gly Ile Gln Pro His Gln Asn Met Ala Leu Phe Arg Tyr Leu Glu Asp  
                                   115                                  120                                  125  
 Lys Ile Asn Ala Glu Glu Lys Thr Ala Trp Ile Thr Asn Ala Ile Thr  
                                   130                                  135                                  140  
 Lys Gly Phe Thr Ala Leu Glu Lys Leu Leu Val Ser Cys Ala Gly Lys  
                                   145                                  150                                  155                                  160  
 Tyr Ala Thr Gly Asp Glu Val Tyr Leu Ala Asp Leu Phe Leu Ala Pro  
                                   165                                  170                                  175  
 Gln Ile His Ala Ala Phe Asn Arg Phe His Ile Asn Met Glu Pro Phe  
                                   180                                  185                                  190  
 Pro Thr Leu Ala Arg Phe Tyr Glu Ser Tyr Asn Glu Leu Pro Ala Phe  
                                   195                                  200                                  205  
 Gln Asn Ala Val Pro Glu Lys Gln Pro Asp Thr Pro Ser Thr Ile  
                                   210                                  215                                  220

&lt;210&gt; 120

&lt;211&gt; 1227

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1224)

&lt;223&gt; coding for fumarylacetoacetate hydrolase (FAAH)

&lt;400&gt; 120

atg gcg ttg ctg aag tct ttc atc gat gtt ggc tca gac tcg cac ttc 48  
 Met Ala Leu Leu Lys Ser Phe Ile Asp Val Gly Ser Asp Ser His Phe  
                   1                                  5                                  10                                  15  
 cct atc cag aat ctc cct tat ggt gtc ttc aaa ccg gaa tcg aac tca 96  
 Pro Ile Gln Asn Leu Pro Tyr Gly Val Phe Lys Pro Glu Ser Asn Ser  
                                   20                                  25                                  30  
 act cct cgt cct gcc gtc gct atc ggc gat ttg gtt ctg gac ctc tcc 144  
 Thr Pro Arg Pro Ala Val Ala Ile Gly Asp Leu Val Leu Asp Leu Ser  
                                   35                                  40                                  45  
 gct atc tct gaa gct ggg ctt ttc gat ggt ctg atc ctt aag gac gca 192  
 Ala Ile Ser Glu Ala Gly Leu Phe Asp Gly Leu Ile Leu Lys Asp Ala  
                                   50                                  55                                  60  
 gat tgc ttt ctt cag cct aat ttg aat aag ttc ttg gcc atg gga cgg 240  
 Asp Cys Phe Leu Gln Pro Asn Leu Asn Lys Phe Leu Ala Met Gly Arg  
                                   65                                  70                                  75                                  80

## 125

cct gcg tgg aag gaa gcg cgt tct acg ctg caa aga atc ttg tca ttt	288
Pro Ala Trp Lys Glu Ala Arg Ser Thr Leu Gln Arg Ile Leu Ser Phe	
85 90 95	
ttg tta ttt ggc ttc aag gtt ttg gtt ttg gta tgt ttt cat gca gct	336
Leu Leu Phe Gly Phe Lys Val Leu Val Leu Val Cys Phe His Ala Ala	
100 105 110	
aat gaa cct atc ttg cga gac aat gat gtt ttg agg aga aaa tca ttc	384
Asn Glu Pro Ile Leu Arg Asp Asn Asp Val Leu Arg Arg Lys Ser Phe	
115 120 125	
cat cag atg agt aaa gtg gaa atg att gtt cct atg gtg att ggg gac	432
His Gln Met Ser Lys Val Glu Met Ile Val Pro Met Val Ile Gly Asp	
130 135 140	
tat aca gac ttc ttt gca tct atg cat cac gcg aag aac tgc gga ctt	480
Tyr Thr Asp Phe Phe Ala Ser Met His His Ala Lys Asn Cys Gly Leu	
145 150 155 160	
atg ttc cgt ggg cct gag aat gcg ata aac cca aat tgg ttt cgt ctt	528
Met Phe Arg Gly Pro Glu Asn Ala Ile Asn Pro Asn Trp Phe Arg Leu	
165 170 175	
ccc att gca tat cat gga cgg gca tca tct att gtc atc tct ggg act	576
Pro Ile Ala Tyr His Gly Arg Ala Ser Ser Ile Val Ile Ser Gly Thr	
180 185 190	
gac att att cga cca aga ggt cag ggc cat cca caa gga aac tct gaa	624
Asp Ile Ile Arg Pro Arg Gly Gln Gly His Pro Gln Gly Asn Ser Glu	
195 200 205	
cca tat ttt gga cct tcg aag aaa ctt gat ttt gag ctt gag atg gct	672
Pro Tyr Phe Gly Pro Ser Lys Lys Leu Asp Phe Glu Leu Glu Met Ala	
210 215 220	
gct gtg gtt ggt cca gga aat gaa ttg gga aag cct att gac gtg aat	720
Ala Val Val Gly Pro Gly Asn Glu Leu Gly Lys Pro Ile Asp Val Asn	
225 230 235 240	
aat gca gcc gat cat ata ttt ggt cta tta ctg atg aat gac tgg agt	768
Asn Ala Ala Asp His Ile Phe Gly Leu Leu Leu Met Asn Asp Trp Ser	
245 250 255	
gct agg gat att cag gcg tgg gag tat gta cct ctt ggt cct ttc ctg	816
Ala Arg Asp Ile Gln Ala Trp Glu Tyr Val Pro Leu Gly Pro Phe Leu	
260 265 270	
ggg aag agt ttt ggg act act ata tcc cct tgg att gtt acc ttg gat	864
Gly Lys Ser Phe Gly Thr Thr Ile Ser Pro Trp Ile Val Thr Leu Asp	
275 280 285	
gcg ctt gag cct ttt ggt tgt caa gct ccc aag cag gat cca cct cca	912
Ala Leu Glu Pro Phe Gly Cys Gln Ala Pro Lys Gln Asp Pro Pro Pro	
290 295 300	
ttg cca tat ttg gct gag aaa gag tct gta aat tac gat atc tcc ttg	960
Leu Pro Tyr Leu Ala Glu Lys Glu Ser Val Asn Tyr Asp Ile Ser Leu	
305 310 315 320	
gag cta gca cac cat acc gtt aac ggt tgc aat ttg agg cct ggt gat	1008
Glu Leu Ala His His Thr Val Asn Gly Cys Asn Leu Arg Pro Gly Asp	
325 330 335	
ctc ctt ggc aca gga acc ata agc gga ccg gag cca gat tca tat ggg	1056
Leu Leu Gly Thr Gly Thr Ile Ser Gly Pro Glu Pro Asp Ser Tyr Gly	
340 345 350	

## 126

tgc cta ctt gag ttg aca tgg aat gga cag aaa cct cta tca ctc aat 1104  
 Cys Leu Leu Glu Leu Thr Trp Asn Gly Gln Lys Pro Leu Ser Leu Asn  
                   355                                  360                                  365

gga aca act cag acg ttt ctc gaa gac gga gac caa gtc acc ttc tca 1152  
 Gly Thr Thr Gln Thr Phe Leu Glu Asp Gly Asp Gln Val Thr Phe Ser  
                   370                                  375                                  380

ggt gta tgc aag gga gat ggt tac aat gtt ggg ttt gga aca tgc aca 1200  
 Gly Val Cys Lys Gly Asp Gly Tyr Asn Val Gly Phe Gly Thr Cys Thr  
                   385                                  390                                  395                                  400

ggg aaa att gtt cct tca ccg cct tga 1227  
 Gly Lys Ile Val Pro Ser Pro Pro  
   405

&lt;210&gt; 121

&lt;211&gt; 408

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 121

Met Ala Leu Leu Lys Ser Phe Ile Asp Val Gly Ser Asp Ser His Phe  
   1                                  5                                  10                                  15

Pro Ile Gln Asn Leu Pro Tyr Gly Val Phe Lys Pro Glu Ser Asn Ser  
                                   20                                  25                                  30

Thr Pro Arg Pro Ala Val Ala Ile Gly Asp Leu Val Leu Asp Leu Ser  
                                   35                                  40                                  45

Ala Ile Ser Glu Ala Gly Leu Phe Asp Gly Leu Ile Leu Lys Asp Ala  
                                   50                                  55                                  60

Asp Cys Phe Leu Gln Pro Asn Leu Asn Lys Phe Leu Ala Met Gly Arg  
                                   65                                  70                                  75                                  80

Pro Ala Trp Lys Glu Ala Arg Ser Thr Leu Gln Arg Ile Leu Ser Phe  
                                   85                                  90                                  95

Leu Leu Phe Gly Phe Lys Val Leu Val Leu Val Cys Phe His Ala Ala  
                                   100                                  105                                  110

Asn Glu Pro Ile Leu Arg Asp Asn Asp Val Leu Arg Arg Lys Ser Phe  
                                   115                                  120                                  125

His Gln Met Ser Lys Val Glu Met Ile Val Pro Met Val Ile Gly Asp  
                                   130                                  135                                  140

Tyr Thr Asp Phe Phe Ala Ser Met His His Ala Lys Asn Cys Gly Leu  
                                   145                                  150                                  155                                  160

Met Phe Arg Gly Pro Glu Asn Ala Ile Asn Pro Asn Trp Phe Arg Leu  
                                   165                                  170                                  175

Pro Ile Ala Tyr His Gly Arg Ala Ser Ser Ile Val Ile Ser Gly Thr  
                                   180                                  185                                  190

Asp Ile Ile Arg Pro Arg Gly Gln Gly His Pro Gln Gly Asn Ser Glu  
                                   195                                  200                                  205

Pro Tyr Phe Gly Pro Ser Lys Lys Leu Asp Phe Glu Leu Glu Met Ala  
                                   210                                  215                                  220

Ala Val Val Gly Pro Gly Asn Glu Leu Gly Lys Pro Ile Asp Val Asn  
                                   225                                  230                                  235                                  240

Asn Ala Ala Asp His Ile Phe Gly Leu Leu Leu Met Asn Asp Trp Ser  
                                   245                                  250                                  255

## 127

Ala Arg Asp Ile Gln Ala Trp Glu Tyr Val Pro Leu Gly Pro Phe Leu  
 260 265 270

Gly Lys Ser Phe Gly Thr Thr Ile Ser Pro Trp Ile Val Thr Leu Asp  
 275 280 285

Ala Leu Glu Pro Phe Gly Cys Gln Ala Pro Lys Gln Asp Pro Pro Pro  
 290 295 300

Leu Pro Tyr Leu Ala Glu Lys Glu Ser Val Asn Tyr Asp Ile Ser Leu  
 305 310 315 320

Glu Leu Ala His His Thr Val Asn Gly Cys Asn Leu Arg Pro Gly Asp  
 325 330 335

Leu Leu Gly Thr Gly Thr Ile Ser Gly Pro Glu Pro Asp Ser Tyr Gly  
 340 345 350

Cys Leu Leu Glu Leu Thr Trp Asn Gly Gln Lys Pro Leu Ser Leu Asn  
 355 360 365

Gly Thr Thr Gln Thr Phe Leu Glu Asp Gly Asp Gln Val Thr Phe Ser  
 370 375 380

Gly Val Cys Lys Gly Asp Gly Tyr Asn Val Gly Phe Gly Thr Cys Thr  
 385 390 395 400

Gly Lys Ile Val Pro Ser Pro Pro  
 405

&lt;210&gt; 122

&lt;211&gt; 11667

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: supression  
 construct 2 p3300.1-Toc159-GFP-RNAi

&lt;400&gt; 122

aattcgtttc tccataataa tgtgtgagta gttcccagat aagggaatta ggggttcctat 60  
 agggtttcgc tcatgtgttg agcatataag aaacccttag tatgtatttg tatttgtaaa 120  
 atacttctat caataaaatt tctaattcct aaaacccaaa tccagtacta aaatccagat 180  
 cccccaatt aattcggcgt taattcagca attcgtaatc atgggtcatag ctgtttcctg 240  
 tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgtg 300  
 aagcctgggg tgcctaataa gtgagctaac tcacattaat tgcgttgccg tcaactgccc 360  
 ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 420  
 gaggcggttt gcgtattggc tagagcagct tgccaacatg gtggagcacg acactctcgt 480  
 ctactccaag aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca 540  
 acaaagggtg atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat 600  
 caaaaggaca gtagaaaagg aagggtggc acacaaatgc catcattgcy ataaaggaaa 660  
 ggctatcggt caagatgcct ctgccgacag tgggtcccaa gatggacccc caccacgag 720  
 gagcatcggt gaaaaagaag acgttccaac cagctcttca aagcaagtgg attgatgtga 780  
 taacatgggt gagcacgaca ctctcgtcta ctccaagaat atcaaagata cagtcctcaga 840  
 agaccaaaagg gctattgaga cttttcaaca aagggttaata tcgggaaacc tcctcggatt 900  
 ccattgcccc gctatctgtc acttcatcaa aaggacagta gaaaagggaag gtggcaccta 960  
 caaatggcat cattgcgata aaggaaaggc tatcgttcaa gatgcctctg ccgacagtgg 1020  
 tcccaaagat ggacccccac ccacgaggag catcgtggaa aaagaagacg ttccaaccac 1080  
 gtcttcaaag caagtggatt gatgtgatat ctccactgac gtaagggatg acgcacaatc 1140  
 ccactatcct tcgcaagacc ttctctata taaggaagtt catttcattt ggagaggaca 1200  
 cgctgaaatc accagtctct ctctacaaat ctatctctct cgagtctacc atgagcccag 1260  
 aacgacgccc ggccgacatc cgccgtgcca ccgaggcgga catgccggcg gtctgcacca 1320  
 tcgtcaacca ctacatcgag acaagcacgg tcaacttccg taccgagccg caggaaccgc 1380

```

aggagtggac ggacgacctc gtccgtctgc gggagcgcta tccctggctc gtcgccgagg 1440
tggaaggcga ggtcgccggc atcgccctac cggggcccctg gaaggcacgc aacgcctacg 1500
actggacggc cgagtcgacc gtgtacgtct ccccccgcca ccagcggacg ggactggggt 1560
ccacgctcta caccacctg ctgaagtccc tggaggcaca gggcttcaag agcgtggctg 1620
ctgtcatcgg gctgcccac gaccgagcg tgcgcataca cgaggcgctc ggatatgcc 1680
cccgcggcat gctgcgggcg gccggcttca agcacgggaa ctggcatgac gtgggtttct 1740
ggcagctgga cttcagcctg ccggtaccgc ccggtccggt cctgcccgtc accgagattt 1800
gactcgagtt tctccataat aatgtgtgag tagttcccag ataagggaat tagggttcct 1860
atagggtttc gctcatgtgt tgagcatata agaaaccctt agtatgtatt tgtatttgta 1920
aaatacttct atcaataaaa tttctaattc ctaaaaccaa aatccagtac taaaatccag 1980
atcccccgaa ttaattcggc gttaattcag tacattaaaa acgtccgcaa tgtgttatta 2040
agttgtctaa gcgtcaattt gtttacacca caatatatcc tgccaccagc cagccaacag 2100
ctccccgacc ggcagctcgg cacaaaatca ccactcgata caggcagccc atcagtcagg 2160
gacggcgta cggggagagc cgttgtaagg cggcagactt tgctcatgtt accgatgcta 2220
ttcggaagaa cggcaactaa gctgccgggt ttgaaacacg gatgatctcg cggagggttag 2280
catgttgatt gtaacgatga cagagcggtg ctgcctgtga tcaccgcggt ttcaaatcgc 2340
gtcccgctga tactatgtta tacgccaact ttgaaaacaa ctttgaaaaa gctgttttct 2400
ggtatttaag gttttagaat gcaaggaaca gtgaattgga gttcgtcttg ttataattag 2460
cttcttgggg tatctttaa tactgtagaa aagaggaagg aaataataaa tggctaaaat 2520
gagaatatca ccggaattga aaaaactgat cgaaaaatac cgctgcgtaa aagatacgga 2580
aggaatgtct cctgctaagg tatataagct ggtgggagaa aatgaaaacc tatatttaa 2640
aatgacggac agccggtata aagggaccac ctatgatgtg gaacgggaaa aggacatgat 2700
gctatggctg gaaggaaagc tgccgtgtcc aaaggtcctg cactttgaac ggcgatgtgg 2760
ctggagcaat ctgctcatga gtgaggccga tggcgctcct tgctcggag agtatgaaga 2820
tgaacaaagc cctgaaaaga ttatcgagct gtatgaggag tgcatcaggc tctttcactc 2880
catcgacata tcggattgtc cctatacgaa tagcttagac agccgcttag ccgaattgga 2940
ttacttactg aataacgatc tggccgatgt ggattgcgaa aactgggaag aagacactcc 3000
atttaaagat ccgcgcgagc tgtatgattt tttaaagacg gaaaagcccg aagaggaact 3060
tgtcttttcc caccgcgacc tgggagacag caacatcttt gtgaaagatg gcaaagtaag 3120
tggctttatt gatcttggga gaagcggcag ggcggacaag tggatatgac ttgccttctg 3180
cgtccggtcg atcagggagg atatcgggga agaacagtat gtcgagctat tttttgactt 3240
actggggatc aagcctgatt gggagaaaat aaaatattat attttactgg atgaattggt 3300
ttagtaccta gaatgcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc 3360
agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaactcg 3420
ctgcttgcaa acaaaaaaac caccgctacc agcgggtggtt tgtttgcccg atcaagagct 3480
accaactcct tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct 3540
tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct 3600
cgctctgcta atcctgttac cagtggctgc tggcagtggc gataagtcgt gtcttaccgg 3660
gttggactca agacgatagt tacgggataa ggcgcagcgg tcgggctgaa cgggggggtc 3720
gtgcacacag ccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga 3780
gctatgagaa agcgcacgc tcccgaggc gagaaaggcg gacaggtatc cggtaagcgg 3840
cagggtcgga acaggagagc gcacgaggga gcttcaggg ggaaacgcct ggtatcttta 3900
tagtctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg 3960
ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg 4020
ctggcctttt gctcacatgt tctttctgc gttatcccct gattctgtgg ataaccgtat 4080
tacgcctttt gagttagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc 4140
agttagcgag gaagcggaag agcgcctgat gcggtatttt ctcttacgc atctgtcgg 4200
tatttcacac cgcataatgg gactctcag tacaatctgc tctgatgccg catagttaag 4260
ccagtataca ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acaccgcca 4320
acaccgctg acgcgcctg acgggcttgt ctgctcccgg catccgctta cagacaagct 4380
gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg 4440
aggcagggtg ccttgatgtg ggcgcggcg gtcgagtggc gacggcgcgg cttgtccgcg 4500
ccctggtaga ttgcctggcc gtagggcagc catttttgag cggccagcgg ccgcgatagg 4560
ccgacgcgaa gcggcggggc gtagggagcg cagcgaccga agggtaggcg ctttttgag 4620
ctcttcggct gtgcgctggc cagacagtta tgcacaggcc aggcgggttt taagagtttt 4680
aataagtttt aaagagtttt aggcggaaaa atcgcccttt ttctctttta tatcagtcac 4740
ttacatgtgt gaccggttcc caatgtacgg ctttgggttc ccaatgtacg ggttccggtt 4800

```

cccaatgtac	ggcttttgggt	tcccaatgta	cgtgctatcc	acaggaaaga	gaccttttcg	4860
acctttttcc	cctgctaggg	caatttgccc	tagcatctgc	tccgtacatt	aggaaccggc	4920
ggatgcttcg	ccctcgatca	ggttgcggta	gcgcactgact	aggatcgggc	cagcctgccc	4980
cgcctcctcc	ttcaaatcgt	actccggcag	gtcatttgac	ccgatcagct	tgcgacagg	5040
gaaacagaac	ttcttgaact	ctccggcgct	gccactgcgt	tcgtagatcg	tcttgaacaa	5100
ccatctggct	tctgccttgc	ctgcggcgcg	gcgtgccagg	cggtagagaa	aacggccgat	5160
gccgggatcg	atcaaaaagt	aatcgggggtg	aaccgtcagc	acgtccgggt	tcttgccctc	5220
tgtgatctcg	cgggtacatcc	aatcagctag	ctcgatctcg	atgtactccg	gccgcccggg	5280
ttcgctcttt	acgatcttgt	agcggctaata	caaggcttca	ccctcggata	ccgtcaccag	5340
gcggcggttc	ttggccttct	tcgtacgctg	catggcaacg	tgcggtggtg	ttaaccgaat	5400
gcagggttct	accaggctgt	ctttctgctt	tccgccatcg	gctcgccggc	agaacttgag	5460
tacgtccgca	acgtgtggac	ggaacacgcg	gccgggcttg	tctcccttcc	cttcccggta	5520
tcggttcatg	gattcgggta	gatgggaaac	cgccatcagt	accaggctcg	aatcccacac	5580
actggccatg	cgggcgggcc	ctgcggaaac	ctctacgtgc	ccgtctggaa	gctcgtagcg	5640
gatcacctcg	ccagctcgtc	ggtcacgctt	cgacagacgg	aaaacggcca	cgtccatgat	5700
gctgcgacta	tcgcgggtgc	ccacgtcata	gagcatcgga	acgaaaaaat	ctggttgctc	5760
gtcgcccttg	ggcggtctcc	taatcgacgg	cgcacgggct	gccggcggtt	gccgggattc	5820
tttgcggtt	cgatcagcgg	ccgcttgcca	cgattcacgg	ggcggtgctt	ctgcctcgat	5880
gcgttgccgc	tgggcggcct	gcgcggcctt	caacttctcc	accaggctat	caccagcgcc	5940
cgcgcgatt	tgtaccgggc	cggatgggtt	gcgaccgtca	cgccgattcc	tcgggcttgg	6000
gggttccagt	gccattgcag	ggccggcaga	caaccagcc	gcttacgcct	ggccaaccgc	6060
ccgttccctc	acacatgggg	cattccacgg	cgtcggtgcc	tggttggtct	tgattttcca	6120
tgcgcctcc	tttagccgct	aaaattcate	tactcattta	ttcatttgct	catttactct	6180
ggtagctgcg	cgatgtattc	agatagcagc	tcggtaatgg	tcttgccctg	gcgtaccgcg	6240
tacatcttca	gcttggtgtg	atcctccgcc	ggcaactgaa	agttgaccgg	cttcatggct	6300
ggcgtgtctg	ccaggctggc	caacgttgca	gccttgctgc	tgcggtgcgt	cggacggccg	6360
gcacttagcg	tggttggtgt	tttgctcatt	ttctctttac	ctcattaact	caaagtgtt	6420
ttgatttaat	ttcagcggcc	agcgccgtga	cctcgcgggc	agcgctcgcc	tcgggttctg	6480
attcaagaac	ggttgtgccc	gcggcgccag	tgccgtgggt	gctcacgcgc	tgcgtagatac	6540
gggactcaag	aatgggcagc	tcgtaccggg	ccagcgctc	ggcaacctca	ccgccgatgc	6600
gcgtgccttt	gatcgcccg	gacacgacaa	aggccgcttg	tagccttcca	tccgtgacct	6660
caatgcgctg	cttaaccagc	tccaccagg	cggcggtggc	ccatatgtcg	taagggtctg	6720
gctgcaccgg	aatcagcag	aagtgcggctg	ccttgatcgc	ggacacagcc	aagtccggcg	6780
cctggggcgc	tccgtcgatc	actacgaagt	cgcgcgggcc	gatggccttc	acgtcgcggt	6840
caatcgtcgg	gcgggtcgatg	ccgacaacgg	ttagcgggtg	atcttcccg	acggccgccc	6900
aatcgcgggc	actgccttgg	ggatcggaat	cgactaacag	aacatcgcc	ccggcgagtt	6960
gcagggcgcg	ggctagatgg	gttgcgatgg	tcgtcttgcc	tgaccgcct	ttctggttaa	7020
gtacagcgat	aaccttcatg	cgttccctt	gcgtatttgt	ttatttactc	atcgcatcat	7080
atacgcagcg	accgcatgac	gcaagctgtt	ttactcaaat	acacatcacc	tttttagacg	7140
gcggcgctcg	gtttcttcag	cggccaagct	ggccggccag	gccgccagct	tggcatcaga	7200
caaaccggcc	aggatttcat	gcagccgcac	ggttgagacg	tgcgcgggcg	gctcgaacac	7260
gtaccgggcc	gcgatcatct	ccgcctcgat	ctcttcggta	atgaaaaacg	gttcgtcctg	7320
gccgtcctgg	tgcggtttca	tgcttggttc	tcttgcggtt	cattctcggc	ggccgccagg	7380
gcgtcgccct	cggctcaatgc	gtcctcacgg	aaggcacgcg	gccgcctggc	ctcggtgggc	7440
gtcacttctc	cgctgcgctc	aagtgcgcgg	tacagggtcg	agcgatgcac	gccaagcagt	7500
gcagccgcct	ctttcacggg	gcggccttcc	tggtcgatca	gctcgcgggc	gtgcgcgac	7560
tgtgcggggg	tgagggtagg	gcgggggcca	aacttcacgc	ctcgggcctt	ggcggcctcg	7620
cgcgcgctcc	gggtgcggtc	gatgattagg	gaacgctcga	actcggaat	gccggcgaac	7680
acggtaaca	ccatgcggcc	ggccggcggtg	gtggtgtcgg	cccacggctc	tgccaggcta	7740
cgcaggcccg	cgcgggcctc	ctggatgcgc	tcggcaatgt	ccagtaggtc	gcgggtgctg	7800
cgggccaggc	ggtctagcct	ggtcactgtc	acaacgtcgc	cagggcgtag	gtggtcaagc	7860
atcctggcca	gctcggggcg	gtcgcgctg	gtgcgggtga	tcttctcgga	aaacagcttg	7920
gtgcagccgg	ccgcgtgcag	ttcgggccgt	tggttggtca	agtctggtc	gtcggtgctg	7980
acgcgggcat	agcccgagcag	gccagcgggc	gcgctcttgt	tcatggcgta	atgtctccgg	8040
ttctagtcgc	aagtattcta	ctttatgcga	ctaaaacacg	cgacaagaaa	acgccaggaa	8100
aagggcaggg	cggcagcctg	tcgcgtaact	taggacttgt	gcgacatgtc	gttttcagaa	8160
gacggctgca	ctgaacgtca	gaagccgact	gcactatagc	agcggagggg	ttggatcaaa	8220



gtactttgat	cccgagggga	accctgtggt	tggcatgcac	atacaaatgg	acgaacggat	8280
aaaccttttc	acgccctttt	aaatatccgt	tatttctaata	aacgctcttt	tctcttaggt	8340
ttacccgccca	atatabcctg	tcaaacactg	atagttttaa	ctgaaggcgg	gaaacgcaca	8400
tctgatccaa	gctcaagctg	ctctagcatt	cgccattcag	gctgcgcaac	tgttgggaag	8460
ggcgatcggg	gcgggcctct	tcgctattac	gccagctggc	gaaaggggga	tgtgctgcaa	8520
ggcgattaag	ttgggtaacg	ccagggtttt	cccagtcacg	acgttgtaaa	acgacggcca	8580
gtgccaaagct	tttggttaga	gcagcttgcc	aacatggtgg	agcacgacac	tctcgtctac	8640
tccaagaata	tcaaagatac	agtctcagaa	gaccaaaggg	ctattgagac	ttttcaacaa	8700
agggtaatat	cgggaaacct	cctcggattc	cattgcccag	ctatctgtca	cttcatcaaa	8760
aggacagtag	aaaaggaagg	tggcacctac	aaatgccatc	attgcgataa	aggaaaggct	8820
atcgttcaag	atgcctctgc	cgacagtggg	cccaaagatg	gacccccacc	cacgaggagc	8880
atcgtggaaa	aagaagacgt	tccaaccacg	tcttcaaagc	aagtggattg	atgtgataac	8940
atggtggagc	acgacactct	cgtctactcc	aagaatatca	aagatacagt	ctcagaagac	9000
caaagggcta	ttgagacttt	tcaacaaagg	gtaatatcgg	gaaacctcct	cggattccat	9060
tgcccagcta	tctgtcactt	catcaaaagg	acagtagaaa	aggaagggtg	cacctacaaa	9120
tgccatcatt	gcgataaagg	aaaggctatc	gttcaagatg	cctctgccga	cagtgggtccc	9180
aaagatggac	ccccacccac	gaggagcatc	gtggaaaaag	aagacgttcc	aaccacgtct	9240
tcaaagcaag	tggtattgatg	tgatatctcc	actgacgtaa	gggatgacgc	acaatcccac	9300
tatccttcgc	aagaccttcc	tctatataag	gaagttcatt	tcatttggag	aggacacgct	9360
gaaatcacca	gtctctctct	acaaatctat	ctctccatgg	catgttctgc	aggtcgactc	9420
tagaggatcc	ccgggtaccg	agctcgaaga	tcttcgacgt	cgggaattcat	ggactcaaag	9480
tcggttactc	cagaaccaac	caaccccttc	tacgttctct	cggggcaatc	aggaaaaacc	9540
tatgcttctg	ttgtcgccgc	cgctgctgct	gcagccgcgc	ataaggagga	tgggtggtgct	9600
gtgagtagtg	ccaaggagtt	ggattcctca	tcggaggctg	tgtctggtaa	ttcggataag	9660
gttggagctg	atgatttata	tgactccgag	aaggagaagc	cgaatttggt	gggtgatggg	9720
aaggtttccg	acgaggtgga	tggttcttta	aaggaggatt	ctactactcc	tgaggctact	9780
ccgaagcctg	agggtggttc	tgggtgagaca	attgggtgtag	atgatgtttc	atcgttatct	9840
ccgaagccgg	aggctgtttc	tgatggtgta	ggggttgtgg	aggagaataa	gaaggtttaag	9900
gaggacgtgg	aggatattaa	agacgatggg	gagagtaaga	ttgaaaatgg	gagtgttgat	9960
gttgatgtga	aacaggcttc	cacagatggg	gagagtga	aagcttccaa	cacttgtcac	10020
tactttctct	tatggtgttc	aatgcttttc	aagataccca	gatcatatga	aacggcatga	10080
cttcttcaag	agcgccatgc	ctgagggata	cgtgcaggag	aggaccatct	tcttcaagga	10140
cgacgggaac	tacaagacac	gtgctgaagt	caagtttgag	ggagacaccc	tcgtcaacag	10200
gatcgagctt	aagggaatcg	atttcaagga	ggacggaaac	atcctcggcc	acaagttgga	10260
atacaactac	aactcccaca	acgtatacat	catggccgac	aagcaaaaga	acggcatcaa	10320
agccaacttc	aagaccgcgc	acaacatcga	agacggcggc	gtgcaactcg	ctgatcatta	10380
tcaacaaaat	actccaattg	gcgatggccc	tgtcctttta	ccagacaacc	attacctgtc	10440
cacacaatct	gccctttcga	aagatcccac	cgaaaagaga	gaccacatgg	tccttcttga	10500
gtttgtaaca	gctgctggga	ttacacatgg	catggatgaa	ctatacaaac	atgatgagct	10560
ttaaggatcc	ttaaagctca	tcagtgttgt	atagttcatc	catgccatgt	gtaatcccag	10620
cagctgttac	aaactcaaga	aggaccatgt	ggtctctctt	ttcgggtggga	tctttcgaaa	10680
gggcagattg	tgtggacagg	taatggttgt	ctggtaaaag	gacagggcca	tcgccaattg	10740
gagtattttg	ttgataatga	tcagcgagtt	gcacgcgcgc	gtcttcgatg	ttgtggcggg	10800
tcttgaagtt	ggctttgatg	ccgttctttt	gcttgtcggc	catgatgtat	acgttgtggg	10860
agttgtagtt	gtattccaac	ttgtggccga	ggatgtttcc	gtcctccttg	aaatcgatcc	10920
ccttaagctc	gatcctgttg	acgaggggtg	ctccctcaaa	cttgacttca	gcacgtgtct	10980
tgtagttccc	gtcgtccttg	aagaagatgg	tcctctcctg	cacgtatccc	tcaggcatgg	11040
cgctcttgaa	gaagtcatgc	cgtttcatat	gatctgggta	tcttgaaaag	cattgaacac	11100
cataagagaa	agtagtgaca	agtgttgga	gctttctcac	tctcccatc	tgtggaagcc	11160
tgtttcacat	caacatcaac	actcccattt	tcaatcttac	tctcaccatc	gtctttaata	11220
tcctccacgt	cctccttaac	cttcttatcc	tcctccacaa	cccctacacc	atcagaaaca	11280
gcctccggct	tcggagataa	cgatgaaaca	tcactacac	caattgtctc	accagaaacc	11340
acctcaggct	tcggagtagc	ctcaggagta	gtagaatcct	cctttaaaga	accatccacc	11400
tcgtcggaaa	ccttcccac	acccaccaa	ttcggcttct	ccttctcgga	gtcagataaa	11460
tcacagctc	caaccttate	cgaattacca	gacacagcct	ccgatgagga	atccaactcc	11520
ttggcactac	tcacagcacc	accatcctcc	ttatcggcgg	ctgcagcagc	agcggcggcg	11580

## 131

acaacagaag cataggtttt tctgattgc cccgaagaag cgtagaaggg gttggttggt 11640  
tctggagtaa ccgactttga gtccatg 11667

<210> 123

<211> 36

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 123

ctcgaggaat tcatggactc aaagtcgggt actcca 36

<210> 124

<211> 40

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 124

ggatccataa gcaagctttc tctctctccc catctgtgga 40

<210> 125

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 125

aagcttccaa cacttggtcac tacttt 26

<210> 126

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 126

ggatccttaa agctcatcat gtttgt 26